Planetary Atmospheres: Astrobiologically Relevant Icy Worlds and Earth as a Proxy Exoplanet

Thesis by

Siteng Fan

In Partial Fulfillment of the Requirements for

the degree of

Doctor of Philosophy

CALIFORNIA INSTITUTE OF TECHNOLOGY

Pasadena, California

2021

(Defended November 19th, 2020)

Siteng Fan
ORCID: 0000-0002-3041-4680

ACKNOWLEDGEMENTS

It is hard to believe that it has been more than five years since I flew over the Pacific Ocean and came to Caltech. These five years are among the most remarkable times in my life. Doing research in a different school, field, and country is never easy, and I have occasionally wondered whether I could make it. Fortunately, with the help and support from people that I would like to acknowledge below, I succeeded in overcoming all the struggles and tough times. They turned my graduate study into an extraordinary and incredible experience which I will cherish forever.

First of all, I would like to express my most sincere gratitude to my PhD advisor Yuk L. Yung. This thesis would not have been possible without him. I appreciate him teaching me the philosophy of doing research, in addition to scientific knowledge and technical methods. His own life trajectory and everyday life have been inspiring me to carry on my academic career. Yuk is also more than a research advisor. He cares about students, and brought me onto the track when I moved to a new country and started a new type of life in graduate school.

I would like to thank my committee members: Andrew P. Ingersoll, Heather A. Knutson, and Konstantin Batygin. Andy taught me about how to think about and discover scientific questions, and also how to present research to the audience. Heather provided me with considerably useful advice for designing my academic path and my future career. Konstantin showed me a way of doing science and enjoying life that I had never imagined, in addition to being the adviser of one of my projects during my first two years.

I benefited greatly from current and past members in Yuk's research group, the Yuk Army and its extended members. I thank Run-Lie Shia, who is capable and helped me to solve any problems I encountered in research, and Sally Newman, who gave me a number of detailed suggestions about scientific writing. I thank Cheng Li for helping me solve many technical problems in research and how to physically survive, Peter Gao for helping me with the Pluto project and emotional support of dealing with difficulties in everyday research, Xi Zhang for advising me on research on planetary atmospheres and guiding me with the plan of my future

academic career, and Qiong Zhang for first helping me land at Caltech and survive in a new country, as well as persuading me to focus on doing science before he went to J.P. Morgan. It was my privilege to work with group members, Vijay Natraj, Chao Liu, Mao-Chang Liang, King-Fai Li, Zhao-Cheng Zeng, Stuart J. Bartlett, and Yuan Wang, from whom I obtained a lot of help for my research, as well as the students in the Yuk Army, Michael L. Wong, Pushkar Kopparla, Jiazheng Li, Tianhao Le, Yangcheng Luo, and Danica J. Adams. I also would like to thank the students that I supervised, Lixiang Gu, Daniel Zhao, and Linfeng Wan, who successfully brought my ideas into reality. I am also grateful to the scientists in the planetary science community, David Crisp, G. Randall Gladstone, L. A. Young, Donald E. Shemansky, Mark Swain, Jonathan H. Jiang, Joshua A. Kammer, without whom my research could never be done like this.

I thank the support from the GPS division and the planetary science option staffs, including Margaret Carlos, Ulrika Terrones, Julie Lee, and Irma Black, for helping me with administration issues. I could not have focused on my research as I did without their support. My thanks also go to Scott A. Dungan, Michael L. Black, and Jian X. Ou for solving computer problems. I also thank Laura Flower Kim and Daniel Yoder at ISP, who have made my life much easier when I moved to this new country.

I am grateful for my friends in planetary science, especially my office mates when I firstly arrived here: Nancy H. Thomas, Joseph G. O'Rourke, Ana Helena de Oliveira Lobo, and Elizabeth Bailey. They all have left me a first and wonderful impression of what graduate school is like in planetary science. Nancy is always enthusiastic to provide help in my everyday life, and became my trip advisor which allowed me to have visited more than half the number of states in the US. Joe showed me how to achieve success at the same time in aspects more than research, and has trained me to become a pilot, one of my childhood dreams besides becoming a scientist. I have also learned a lot from Ana through things we shared in common: being the only international students in the same option and the same year, having a close research field of atmospheres, and taking the same classes in the same term. I also appreciate Lu Pan, Henry Ngo, Nathaniel T. Stein, Nicole L. Wallack, Stephen Markham, and Shreyas Vissapragada for providing such a warm research environment to

have spent my past five years in.

I cherish the moments spent with my friends at Caltech that have lightened up my life. Special thanks to Erya Yu, Xunyi Wu, Xinying Ren, and Leiya Ma for making my life in graduate school fascinating. I thank Hao Xie, Minyan Zhong, Weimeng Kong, Yanzhe Zhu, and Yuanlong Huang for regular reunions that have relieved my homesickness and for taking me bouldering and letting me realize what a student in the GPS division should be like. I would like to thank Sheng Wang, Xiaozhe Ding, Ronghui Zhu, Yitong Ma for discussions about science in other fields and for keeping reminding me what the 21th century is of. I also thank Jize Yu, Xiao Tong, Zili Fan, Xiaomin Li, Liuchi Li, Xinyan Liu, Yicheng Luo, Shiyu Gu, Yalu Chen, Yang Gao, Ke Yu, Yu Su, Jinglin Huang, Jun Shi, Zijie Qu, and Chujun Lin for routine Friday night board games that have kept me productive in research and frequent sports of badminton and skiing that have allowed me to stay fit. I also thank Zhicai Zhang and De Huang, the first two same-year friends I made, for bringing me to this lovely community.

I would also like to thank my undergraduate advisor in Peking University, Jiansen He, who helped me initiate my research career, and Hui Tian, who provided me with help for applying for graduate schools and encouraged me to pursue a PhD degree in planetary science at Caltech.

Finally, I devote my thesis to my parents, Jun Fan and Hanyi Huang for raising me over the past more than twenty years. I would never have come this far without their trust, love, support, and the best environment to have grown up.

ABSTRACT

"How did we get here?" is a long-standing question in planetary science. Characterizing the pre-biotic atmospheric environment in which life may emerge is critical and increasingly urgent. Given the fact that the Earth provides the only ground truth of habitable worlds, most of the characterizations are based on the current Earth. However, life did not emerge on the modern Earth. It instead took place in a prebiotic environment, which includes a nitrogen-dominated, methane-abundant and oxygen-negligible reducing atmosphere. Therefore, this type of planetary atmospheres has great significance in the context of astrobiology and the search for life. Despite that real time observations can not be obtained for Early-Earth, spacecraft observations of the atmospheres of two icy worlds in the solar system, Titan and Pluto, can provide such valuable constraints. The theme of Chapter 2 and 3 of this thesis focus on this topic of investigating the atmospheres of Titan and Pluto using spectroscopic analysis. Chapter 4 studies the search for life also across the spectrum from a prospective the other way. It characterizes the Earth, the only known inhabited planet, as an exoplanet proxy, to derive observational benchmarks for habitability assessment.

Chapter 2 studies Titan. It retrieves the hydrocarbon and nitrile species in Titan's upper atmosphere using stellar occultation observations obtained by Cassini UltraViolet Imaging Spectrograph (UVIS) during its Titan flybys. An innovative method is introduced to consider the pointing issue of the instrument, which prevents most of the previous spectral analyses. Combing an instrument simulator for handling the pointing motion and the Markov-Chain Monte Carlo (MCMC) method for parameter searching, species abundances in Titan's atmosphere are successfully retrieved during occultations with large pointing motions. The method also obtains the altitude range where the abundance of each species could be constrained.

Chapter 3 studies Pluto. It investigates the morphology and microphysical processes of Pluto's haze particles in the lower 50km of its atmosphere using observations obtained by multiple instruments onboard the New Horizons spacecraft during its Pluto flyby in 2015. It suggests that Pluto's haze particles have a bimodal distribution: a large-size population of

~1μm fractal aggregates, which consists of ~20nm monomers, and a small-size one of ~80nm. This result successfully addresses the disagreement among the instruments, and provides important constraints on transport and dimensional transition of haze particles in Pluto's atmosphere.

Chapter 4 studies exoplanets. It evaluates the observational baseline for Earth-like exoplanets using the Earth as a proxy. Observations of the Earth's images obtained by the Deep Space Climate ObserVatoRy (DSCOVR) are integrated to one single point to generate light curves of the "proxy" planet. Using the singular value decomposition (SVD) method, we found that the surface information of the "proxy" planet is in the second principal component (PC) of its light curves, while the first PC mainly consists of that of clouds. Using the strong linear correlation between the time series of the second PC and the corresponding land fraction, we constructed the first two-dimensional surface map of the Earth seen from a hypothetical distant observer, an observer who treats the Earth as an exoplanet.

PUBLISHED CONTENT AND CONTRIBUTIONS

Fan, et al. (2019). “Retrievals of Abundances of Hydrocarbons and Nitrile Species in Titan’s upper Atmosphere”. In: *Earth and Space Science* 6, pp. 1057–1066. doi: 10.1029/2018EA000477.

S. F. participated in the conception of the project, analyzed the data from instrument, processed the navigation data, developed and tested the retrieval algorithm, and participated in the writing of the manuscript.

-Adapted for the contents in Chapter 2

Fan, et al. (2020). “A bimodal distribution of haze in Pluto's atmosphere”. In preparation.

S. F. participated in the conception of the project, analyzed and calibrated the data, developed and tested the forward scattering model and retrieval algorithm, and participated in the writing of the manuscript.

-Adapted for the contents in Chapter 3

Fan, et al. (2019). “Earth as an Exoplanet: A Two-dimensional Alien Map”. In: *Astrophysical Journal Letters* 882, L1. doi: 10.3847/2041-8213/ab3a49.

S. F. participated in the conception of the project, analyzed the data, developed and tested the decomposition algorithm, constructed the 2D surface map, and participated in the writing of the manuscript.

-Adapted for the contents in Chapter 4

TABLE OF CONTENTS

LIST OF ILLUSTRATIONS

Number *Page*

LIST OF TABLES

Number *Page*

Chapter 1

Introduction

"Are we alone?" is a fundamental question in planetary science. It was raised on the first day of this research field, and still remains as an open question. The search for life beyond Earth is among one of the ultimate goals for decades, which inspired a number of missions exploring solar system bodies and looking for signs of life in other planetary systems. After a decade of extensively discovering and characterizing exoplanet systems, more than four thousand exoplanets have been confirmed to date. Meanwhile, our knowledge on planetary system formation has undergone an explosive growth, which leads to a new era of astrobiology. For the next step, detecting biosignatures on exoplanets has become one of the primary goals of missions in the near future, e.g., James Webb Space Telescope (JWST), Habitable Exoplanet Imaging Mission (HabEx), and Large Ultraviolet Optical Infrared Surveyor (LUVOIR). Given the fact that the Earth provides the only ground truth of habitable worlds (Seager & Bains 2015), most biosignatures are proposed based on the current Earth (Schwieterman et al. 2018). However, life emerged in a prebiotic environment instead of modern Earth. The composition of the Earth's atmosphere has evolved significantly through geological time, and Earth's current atmosphere is representative of only ~13% of the total inhabited history (Figure 1.1). Due to the critical role of the prebiotic atmosphere, environments resembling the early Earth have greater significance in the context of astrobiology and the search for life.

The early Earth is thought to have had a nitrogen-dominated reducing atmosphere, which contained abundant methane and negligible oxygen (Catling & Zahnle 2020). Solar UV photons initiated photochemistry and aerosol formation, which are both important to the emergence of life. The prebiotic photochemistry is important because it provided a pathway for small organic molecules to grow. Products of these chemical processes, especially the organic molecules with aromatic rings, provided the life building materials. Aerosol particles, macromolecular organic molecules, are also among the important products of chemistry and microphysics. Photochemistry was initiated by methane photolysis at the top

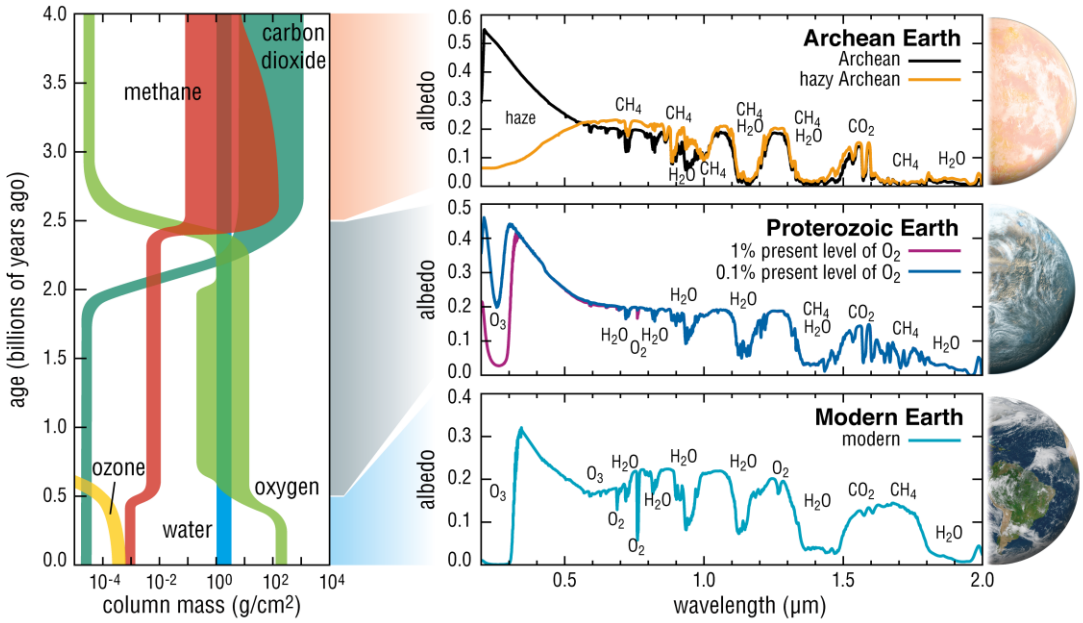

Figure 1.1. Evolution of Earth's atmosphere through geological time. (Credit: Giada Arney)

of the atmosphere. When transported downwards, the chemical products became larger and larger and finally formed aerosol particles through nucleation and coagulation. Aerosol is critical in providing a suitable environment for life emergence. It significantly influenced the energy budget of the Early-Earth through radiation, which results in a suitable temperature of Early-Earth, and also had UV shielding effect that avoid bio-harmful solar UV photons reaching the surface of the planet.

Given the fact that it is impossible to obtain real-time observations of Paleo-Earth, analyses of the prebiotic environments are based on proxy measurements. Fortunately, a few solar system icy bodies in the outer solar system currently have similar nitrogen-dominated and methane-abundant reducing atmospheres. Among then, Pluto and Titan can provide fruitful and valuable constraints for prebiotic-like environments with spacecraft observations, which could significantly improve our understanding of the chemical and physical processes in prebiotic environments, and therefore the assessment of exoplanet habitability. Characterizing the atmospheres of Titan and Pluto is the subject of Chapter 2 and 3.

Titan, the second largest moon in the solar system, has a substantial nitrogen-dominated

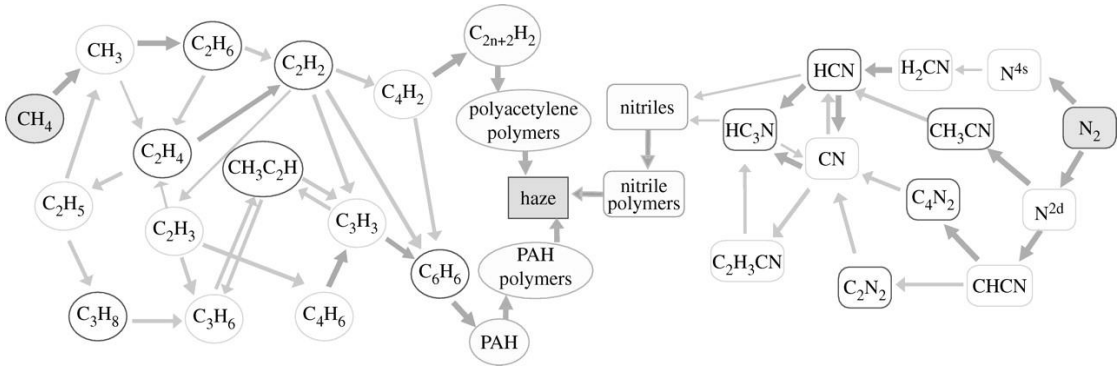

Figure 1.2. Photochemical scheme in Titan's atmosphere. Taken from Atreya et al. (2006).

atmosphere with a surface pressure of ~1.5bar (Lindal et al. 1983). It includes methane at a few percent, which results in a natural laboratory for extraordinary organic chemical processes (Figure 1.2, Atreya et al. 2006), and therefore has significant prebiotic interest. At the top of Titan's atmosphere, incoming solar UV flux dissociates methane molecules to form methyl radicals and excites nitrogen molecules. The small molecules react with each other to form large hydrocarbon and nitrile polymers, including chains and rings, and finally form haze particles in the atmosphere. To investigate and to constrain these chemical processes, observations of the chemical species are required. In early 1980s, the Voyager 1 and 2 missions confirmed the presence of organic chemical species larger than methane (C_2H_2, C_2H_4, C_2H_6, C_3H_8, CH_3C_2H, C_4H_2, etc., Hanel et al. 1981, Kunde et al., 1981). Vertical profiles of the small hydrocarbon and nitrile species had not been obtained until the Cassini-Huygens mission in 2004 (Shemansky et al. 2005). Due to the spectral features of these chemical species, far-ultraviolet (FUV) is an ideal wavelength range for remote sensing to probe the chemical-interesting range of Titan's atmosphere (400-1200km, Figure 1.3, Hörst 2017), as photons within this range are energetic enough to dissociate the organic molecules and therefore show spectral extinction features. Using observations during tens of Cassini's flybys of Titan, seasonal variation of atmospheric dynamics and chemical change could be analyzed. However, due to a technical issue of instrument pointing motion, most of these observations cannot be analyzed until recently. Chapter 2 of this thesis proposes a new methodology of addressing this issue, and makes it feasible to study seasonal cycles of the middle and upper atmosphere of Titan.

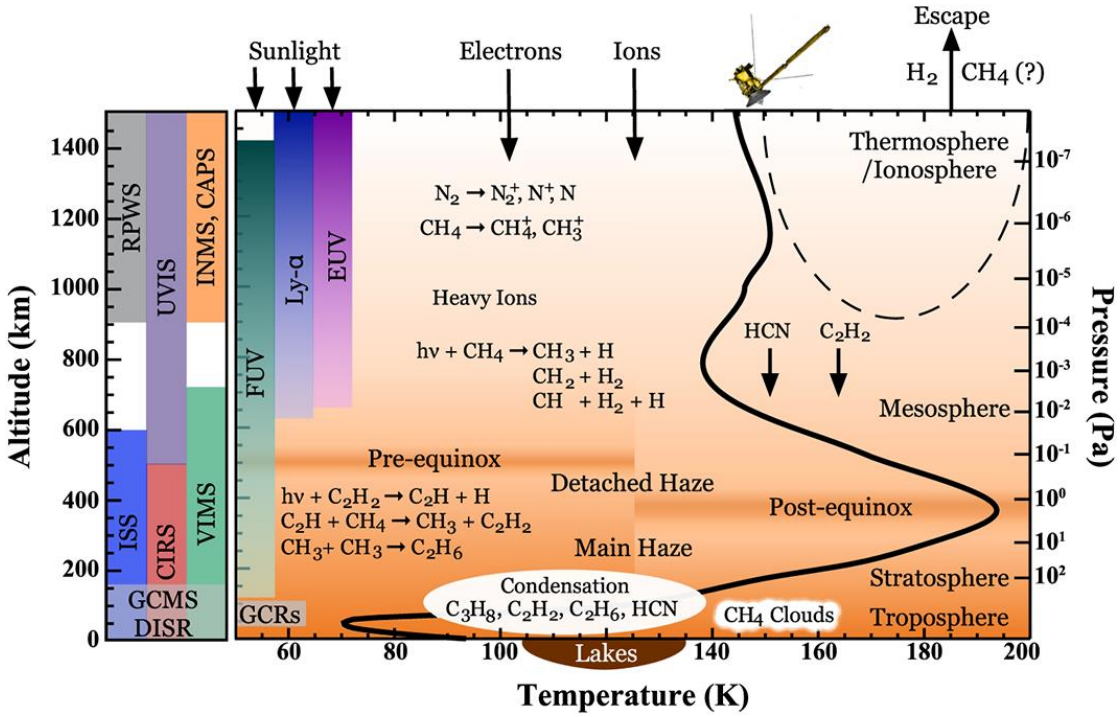

Figure 1.3. Representative temperature profile of Titan's atmosphere, some of the major chemical processes, and the approximate altitude coverage of the instruments carried by Cassini-Huygens. Taken from Hörst (2017).

The existence of Pluto's atmosphere was first confirmed using stellar occultation in late 1980s (Elliot et al. 1989). Due to the large distance between Pluto and Earth, vertically resolvable observations of Pluto's atmosphere were not available until the New Horizons' flyby of Pluto in 2015 (Stern et al. 2015). The atmosphere of Pluto is thin and cold with a surface pressure of ~1Pa when the flyby took place (Gladstone et al. 2016). Nitrogen is at gas-solid equilibrium at the interface of Pluto's atmosphere and surface, so the surface pressure may vary orders of magnitude due to the change of solar insolation. Pluto's atmosphere has similar chemical compositions as that of Titan, which is also nitrogen-dominated with a few percent of methane (Young et al. 2018). Photochemistry is therefore similar in Pluto's atmosphere, except for gas condensation due to Pluto's low temperature (Wong et al. 2017). Organic haze also forms as a product, which is one of the amazing discoveries made by the New Horizons mission (Cheng et al. 2017). The haze in Pluto's atmosphere is thought to be important to the gas condensation, surface color and Pluto's energy budget (Wong et al. 2017, Grundy et al. 2017, Zhang et al. 2017). However, the haze

morphology still remains as a mystery five years after the flyby due to contradicting interpretations of haze observations obtained by different instruments. Therefore, to address the conflict and answer the fundamental question of Pluto's haze, Chapter 3 conducts a joint retrieval of Pluto's haze and searches for a solution that could include all available observations.

In addition to characterizing prebiotic environments that resemble Early-Earth, studying the environment of the modern Earth, the only confirmed currently inhabited environment, from an exoplanet perspective could provide valuable observation benchmarks for assessing exoplanet habitability, which is as important and become increasingly urgent given the timeline of the near future missions. Analyzing distant observables of the Earth, which is treated as a proxy exoplanet in this case, can provide the answer to the fundamental "are we alone?" from the other way. Given the fact that current telescopes and those in the near future could not resolve exoplanets better than single points, photometric observations will remain as the major approach for a long time. Therefore, observations of single-point Earth when the Earth images are spatially integrated to single pixels have great significance in the benchmark analysis. The idea was first raised and performed in 1990s with a few snapshots of the Earth taken by the Galileo spacecraft (Sagan et al. 1993, Geissler et al. 1995), then followed by two one-day light curves obtained by the Deep Impact spacecraft (Cowan et al. 2009, 2011). These results show that gases in the Earth's atmosphere and surface features could be obtained by distant observers, even when the Earth is not spatially resolved. In Chapter 4, this approach is greatly extended using observations from a recent mission, Deep Space Climate Observatory (DSCOVR, Jiang et al. 2018). It provides benchmark observations with significant details which includes ten wavelength channels and a high time resolution of ~1-2 hours. Analysis of spatial features from single-point light curves and influence of their diurnal and seasonal changes becomes feasible.

Chapter 2

Retrieval of Chemical Abundances in Titan's Upper Atmosphere from Cassini UVIS Observations with Pointing Motion

Siteng Fan[1], Donald E. Shemansky[2], Cheng Li[1], Peter Gao[3],

Linfeng Wan[4], Yuk L. Yung[1,5]

[1]California Institute of Technology, Pasadena, CA 91125.

[2]Space Environment Technologies, Altadena, CA 91101.

[3]University of California, Berkeley, CA 94720.

[4]University of California, Santa Cruz, CA 95064.

[5]Jet Propulsion Laboratory, California Institute of Technology, Pasadena, CA 91109.

Fan, et al. (2019). "Retrievals of Abundances of Hydrocarbons and Nitrile Species in Titan's upper Atmosphere". In: *Earth and Space Science* 6, pp. 1057–1066. doi: 10.1029/2018EA000477.

2.1 Abstract

Cassini/UVIS FUV observations of stellar occultations at Titan are well suited for probing its atmospheric composition and structure. However, due to instrument pointing motion, only five out of tens of observations have been analyzed. We present an innovative retrieval method that corrects for the effect of pointing motion by forward modeling the Cassini/UVIS instrument response function with the pointing motion value obtained from the SPICE C-kernel along the spectral dimension. To illustrate the methodology, an occultation observation made during flyby T52 is analyzed, when the Cassini spacecraft had insufficient attitude control. A high-resolution stellar model and an instrument response simulator that includes the position of the point source on the detector are used for the analysis of the pointing motion. The Markov Chain Monte-Carlo method is used to retrieve the line-of-sight abundance profiles of eleven species (CH_4, C_2H_2, C_2H_4, C_2H_6, C_4H_2, C_6H_6, HCN, C_2N_2, HC_3N, C_6N_2, and haze particles) in the spectral vector fitting process. We obtain tight constraints on all of the species aside from C_2H_6, C_2N_2, and C_6N_2, for which we only retrieved upper limits. This is the first time that the T52 occultation was used to derive abundances of major hydrocarbon and nitrile species in Titan's upper and middle atmosphere, as pointing motion prohibited prior analysis. With this new method, nearly all of the occultations obtained over the entire Cassini mission could yield reliable profiles of atmospheric composition, allowing exploration of Titan's upper atmosphere over seasons, latitudes, and longitudes.

2.2 Introduction

Stellar occultation observations made by the Cassini Ultraviolet Imaging Spectrograph (UVIS, Esposito et al. 2004) are essential for constraining the photochemistry of Titan's upper atmosphere, where hydrocarbon and nitrile species show spectral features in the ultraviolet (Hörst 2017). During the last decade, vertical profiles of these species have been derived from selected Titan flybys (Shemansky et al. 2005, Liang et al. 2007, Koskinen et al. 2011, Kammer et al. 2013), which have increased our understanding of physical and chemical processes in Titan's atmosphere. Far-UV (FUV) observations are especially

important in constraining the abundances of hydrocarbons more complex than methane, as well as nitriles. However, FUV observations from only five (TB, T21, T41i, T41e, and T53) out of the tens of flybys have been used for retrievals to date (Shemansky et al. 2005, Koskinen et al. 2011, Capalbo et al. 2016) due to instrument pointing motion, which causes movement of the stellar image on the UVIS detector. Pointing motion is introduced by the spacecraft attitude control system, which frequently triggers thrusters during occultation observations (Chiang et al. 1993). Deadband of attitude control during stellar occultations is set as ±0.5 mrad (Pilinski & Lee, 2009), comparable to that of a spectral pixel (1.0 mrad × 0.25 mrad), which leads to a shift in the spectral structure. Consequently, the traditional method of analyzing extinction spectra obtained by dividing the nonzero optical depth target spectrum by the extinction-free target spectrum becomes inaccurate. The spectral distortion is non-linear due to UVIS internal instrument scattering. Attempts have been made to determine pointing motions for other instruments, e.g., Cassini/VIMS (Maltagliati et al., 2015), and similar issues arose for similar UV instruments when observing stellar occultations of Titan's atmosphere (e.g. Voyager 1 UVS, Vervack et al. 2004). Using an instrument simulator for forward modeling is essential for evaluating spectra.

Our proposed method uses a Cassini/UVIS simulator for forward modeling that combines the pointing information obtained from the SPICE C-kernel (NASA NAIF) and line-of-sight (LOS) chemical abundances in Titan's atmosphere to simulate spectra. The Markov Chain Monte-Carlo (MCMC) method (Foreman-Mackey et al. 2013) is used as the parameter search tool. Using our proposed method, we derive for the first time vertical profiles of hydrocarbon and nitrile species from flyby T52, which shows significant pointing motion during the entire duration of the observation.

The remainder of this paper is organized as follows: in Section 2.3, the pointing motion during flyby T52 and its effects are demonstrated; detailed methodology is presented in Section 2.4; analysis of synthetic spectrum for method testing is in Section 2.5; spectral analysis results and vertical profiles of retrieved species are given in Section 2.6, together with brief discussions; our conclusions and the implications for applying this method are

discussed in Section 2.7.

2.3 Compensating for pointing motion

The instrument used in this work is the Cassini/UVIS FUV spectrograph. It covers the spectral range between 1115 and 1912 Å (Esposito et al. 2004). We select flyby T52 for demonstrating our new retrieval method, which took place on April 3rd, 2009, with the stellar source α Eri. The occultation observation during this flyby is divided into two segments due to different integration durations. We analyze the second segment (NASA PDS ID: FUV2009_093_23_55) in this work, which has an integration duration of 1.75 s for each image; the ray tangent height of incoming stellar light above Titan's surface during this segment was within the critical range (0-1500 km) for atmospheric characterization. The geometry of this segment is given in Table 2.1, computed from the SPICE C-kernel (NASA NAIF).

The image of α Eri is a point source compared to the pixel size on the slit. Five spatial pixels with indices 25 to 29 were electronically windowed to record the photon counts. Each pixel has a length of 1.0 mrad along the spatial dimension (length from upper left to lower right in Figure 2.1a) and the low resolution slit width was set to 1.5 mrad (width from upper right to lower left in Figure 2.1a). During the stellar occultation, the spacecraft navigation system controls pointing through a stellar/solar referenced 3-axis stabilized platform with a deadband of 0.5 mrad referenced to the UVIS FUV principal axis (Pilinski & Lee, 2009). The thrusters of the spacecraft attitude control system react only at the

Table 2.1. Stellar occultation during T52.

	Start	Near 1000km	Near 500km	End
Time	2009-093 23:55:27	2009-094 00:10:51	2009-094 00:16:51	2009-094 00:27:49
Latitude (°)	32.6	35.6	37.1	39.1
Longitude (°)	328.9	319.5	312.8	292.7
Ray Height (km)	2347	999	502	-290

*Latitude and longitude are computed for the impact parameter (radial vector from body center). It varied by a few degrees during the observation due to spacecraft motion.

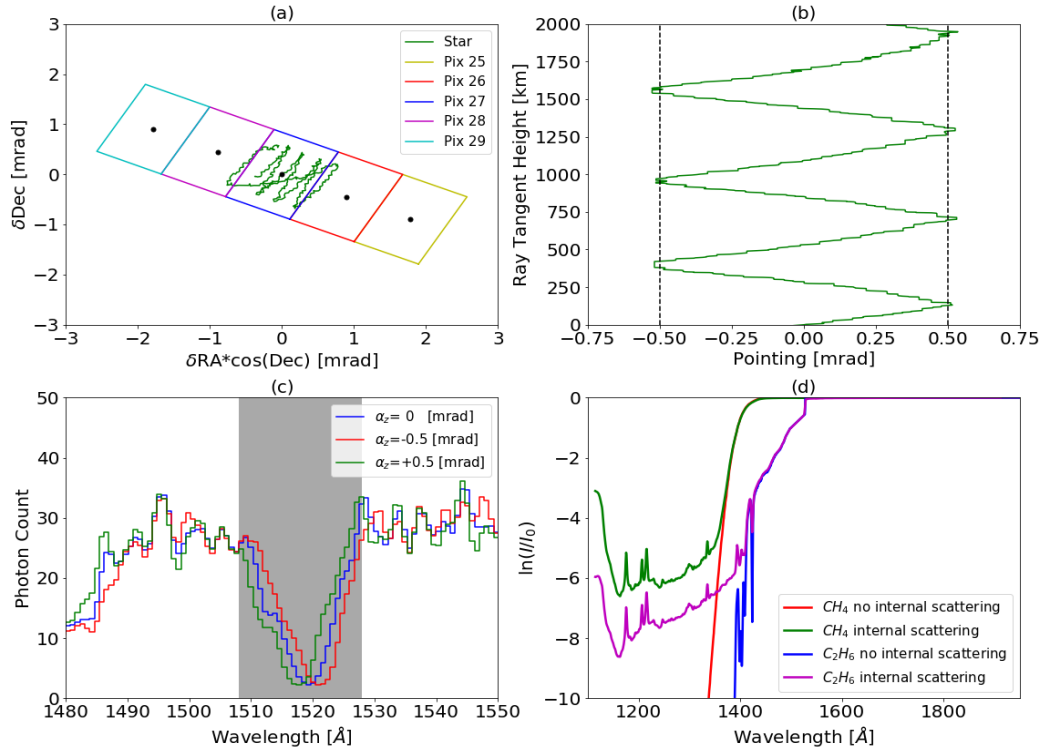

Figure 2.1. Pointing motion and its effects. (a) Spatial orientation of Cassini/UVIS detector pixels (color rectangles) with their centers (black points) and the star motion on the detector (green solid line) obtained from the SPICE C-kernel. The pixel with index 27 (blue rectangle) is the center pixel during the T52 occultation observation. The spatial dimension is along the length of the 5 pixels, with the spectral dimension being perpendicular to that. (b) Vertical profile of pointing motion above Titan's surface along the spectral dimension (solid green line). Deadbands along this dimension at each end are denoted by two vertical black dashed lines. (c) Simulated photon count spectra from the T52 occultation at 754 km ray tangent height with pointing values along the spectral dimension of 0 (blue), -0.5 (red) and +0.5 mrad (green). The shaded area indicates the signature extinction feature of C_2H_2 near 1520 Å. Figure 2.S1 shows the full extent of the three spectra in the FUV wavelength range. (d) Extinction spectra of CH_4 and C_2H_6 with (green for CH_4 and magenta for C_2H_6) and without (blue for CH_4 and red for C_2H_6) taking into account instrument internal scattering. The LOS abundances for both species are set to 10^{18} cm^{-2}, which is approximately the value for that of CH_4 at 750 km.

deadband limit. Therefore, the star image shows motion on the slit (green line in Figure 2.1a), which results in changes in photon count distributions among spatial pixels along the spatial dimension and spectrum distortions along the spectral dimension. The effect of star image motion along the spatial dimension can be eliminated by summing up the photon

counts received by all five spatial pixels after flat fielding. The windowed spatial pixels fully contain the stellar image during the entire T52 occultation. Motion of the star image along the spectral dimension, however, must be modeled to analyze the data, which is the focus of this work.

The star image position on the slit along the spectral dimension as a function of the ray tangent height in Titan's atmosphere during T52 is shown in Figure 2.1b. The star image moved at a constant rate within ±0.5 mrad between the deadband points, showing nominal function of the control system. This introduces strong distortions in the spectral dimension, as demonstrated in Figure 2.1c by the simulated spectra with different pointing motions. An extinction feature of C_2H_2 near 1520 Å is shifted by a few pixels due to the pointing motion, causing a difference in spectral structure between spectra that prevents constructing extinction spectra by dividing one spectrum by another. Moreover, due to the photon scattering effects in the instrument, a single spectral line is spread across a wide range of wavelengths, necessitating simultaneous modeling of a large swath of the spectrum. The full width at half maximum of a typical spectral line is ~1.5 Å, or approximately two spectral pixels, while the line wings extend hundreds of angstroms. This results in non-linearity in the spectrum. Figure 2.1d shows an example of internal photon scattering. Photon counts shortward of 1360 Å are mostly the result of internal scattering from longer wavelengths when CH_4 extinction saturates this range. Therefore, it is not feasible to shift the spectrum back even if the value of the pointing motion were known. A forward instrument model is essential for extracting information from these distorted spectra.

2.4 Methodology

A detailed description of Cassini/UVIS data calibration and reduction is given in Chapter 9 and 10 of the Cassini/UVIS Users Guide, available in NASA-PDS (2017). To improve statistical accuracy, spectra obtained during T52 are integrated over time intervals of 17.5 s, covering an altitude range of ~25 km. The selected integration intervals are merged with pointing information obtained from the SPICE C-kernel (NASA NAIF), which is shown in

Figure 2.1b as the green path. The reference spectrum of the star (I_0) is constructed by averaging the spectra when the ray tangent height is greater than 1500 km above Titan's surface, where extinction by Titan's atmosphere is negligible, and the pointing motion is less than 0.125 mrad (half of a pixel width).

Our forward model combines an extinction model and an instrument simulator. The extinction model computes the intensity spectrum received by Cassini/UVIS, while the instrument simulator (Shemansky et al. 2005; Shemansky & Liu 2012) generates photon count observations based on the intensity spectrum, instrument internal scattering and pointing motion. The instrument simulator contains high resolution response functions for each pixel, which encapsulate the effects of instrument internal scattering that were measured in the lab and calibrated in flight. The core of the point spread function has a full width at half maximum of ~1.7 Å and the wings extend over a spectral range of 800 Å.

Eleven hydrocarbon and nitrile species (CH_4, C_2H_2, C_2H_4, C_2H_6, HCN, C_4H_2, C_6N_2, C_6H_6, haze, HC_3N and C_2N_2), which have extinction features in the FUV, are considered in the forward model. We include two species in the retrieval, C_2N_2 and C_6N_2, which have not been detected previously, to examine the extent to which their abundances can be constrained, as photochemical models suggest their existence (e.g. Willacy et al. 2016). The cross sections of these species are obtained from laboratory work (Table 2.2), some of

Table 2.2. Extinction cross sections.

Species	Reference	Wavelengths (Å)
CH_4	Kameta et al. (2002)	1115-1426
	Chen & Wu (2004)	1426-1490
C_2H_2	Wu et al. (2001)	1115-1912
C_2H_4	Wu et al. (2004)	1115-1912
C_2H_6	Au et al. (1993)	1115-1193
	Wu et al. (2004)	1199-1528
HCN	Nuth & Glicker (1982)	1115-1300
	Lee (1980)	1300-1568
C_4H_2	Ferradaz et al. (2009)	1150-1912
C_6N_2	Connors et al. (1974)	1115-1912
C_6H_6	Pantos et al. (1978)	1115-1912
HC_3N	Ferradaz et al. (2009)	1150-1560
C_2N_2	Nuth & Glicker (1982)	1115-1701

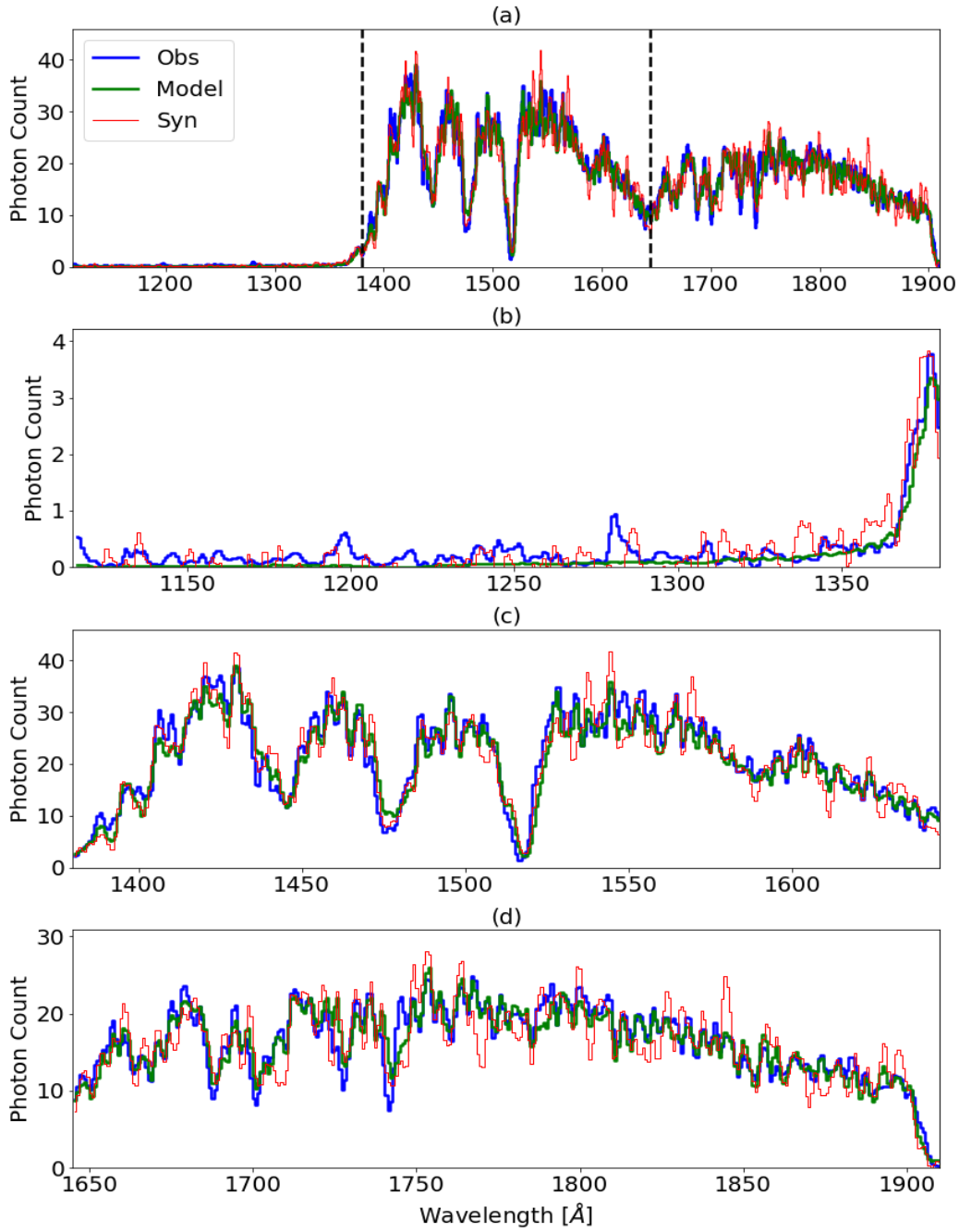

Figure 2.2. (a) Photon count spectra of the T52 occultation at 754 km ray tangent height showing the observed spectrum (blue), the best fit simulated model spectrum (green) and the synthetic spectrum (red), which includes artificially introduced noise. (b)-(d) Detailed views of the photon count spectra split along the black dashed lines in (a). The y axes scales are different in (b)-(d) for the purpose of presentation. The spectral contribution of each species to this spectrum is shown in Figure 2.S2 in the supporting information.

which were conducted at room temperature. The differences in temperatures between that of the measurements and ambient conditions in Titan's atmosphere may contribute ~20% uncertainty to the LOS abundances. Pressure effects are negligible. Haze particles are assumed to be spherical with a radius of 12.5 nm and the same optical properties as their laboratory analogs ("tholin"; Khare et al. 1984), in line with Liang et al. (2007) and Koskinen et al. (2011). We combine the LOS abundances and cross sections to construct spectra based on a normalized I_0.

Retrieval is a multivariable inverse problem. Combining a proper retrieval algorithm with a forward model, physical properties can be derived from the observations. The MCMC method is used in this work to solve the inverse problem. It searches parameter space with the ability to extract asymmetric posterior probability density functions (PDFs) in a computationally feasible way. For each proposed parameter set, a spectrum is constructed with the procedure described above, and is then compared with the observed spectrum to determine the posterior probability of this parameter set. The cost function is defined as follows:

$$\ln(\mathrm{p}) = -\frac{1}{2}\sum_i \left[\frac{\left(I_{Obs_i} - I_{MCMC_i}\right)^2}{\sigma_i^2 + 0.1} + \ln(\sigma_i^2 + 0.1)\right] \qquad (2.1)$$

where p is the posterior probability of one proposed parameter set; I_{Obs} and I_{MCMC} are the photon counts from the observation and those calculated from the forward model during one MCMC attempt, respectively; and σ_i is the standard deviation of the spectral intensity at wavelength i, assumed to be the square root of the simulated photon count. A softening factor of 0.1 is added to the standard deviation at each wavelength to avoid dividing by zero when the intensity decreases to zero at some wavelengths at low altitudes. An example of I_{Obs} and I_{MCMC} is shown in Figure 2.2 as blue and green lines, respectively. With this cost function, we use the emcee package (Foreman-Mackey et al. 2013) to conduct the MCMC parameter search. An MCMC procedure with 120 chains, selected according to the number of parameters, is used to search through parameter space. A bounded uniform prior in log space is set for each parameter, so abundances are retrieved with no prior knowledge

within 2 orders of magnitude of the predicted values from the latest results of the Caltech/JPL photochemical model KINETICS (Li et al. 2014, Li et al., 2015, Willacy et al. 2016). The bounds are adjusted by 2 orders of magnitude if necessary after each 2000 steps according to the PDF of the parameter and the converging criterion. The MCMC procedure is extended for 2000 more steps after the final bound adjustment to generate the resulting PDFs.

2.5 Synthetic spectrum analysis

We analyze a synthetic spectrum (red line in Figure 2.2) to test the reliability of our method. We use the LOS abundances of all species obtained from fitting to the observation (green line in Figure 2.2) as well as the pointing motion (Figure 2.1b) to construct the synthetic spectrum. We then add noise on the level of the square root of the simulated photon count plus the softening factor to obtain the final synthetic spectrum (red line in Figure 2.2) with the same calibration procedure as the data. In other words, the synthetic spectrum is the same as the simulated spectrum used to fit the observations, except it includes noise. Figure 2.2 shows that the noise level is slightly greater than the disagreements between the simulated spectrum and the observation, such that the analysis of the synthetic spectrum should give a lower limit to the reliability of our method.

Figure 2.3 shows the PDFs of the LOS abundances of hydrocarbon and nitrile species resulting from the synthetic spectrum analysis. The LOS abundance values used to construct the synthetic spectrum are shown as black dashed lines. The constraint on each LOS abundance is interpreted by fitting the PDFs with three types of functions, Gaussian, sigmoid, and constant, and the PDFs are categorized as such by comparing the residuals of the fit. LOS abundances with Gaussian-like PDFs are defined as well constrained (e.g., CH_4, C_2H_2, C_2H_4, HCN, C_4H_2, C_6H_6 and HC_3N in Figure 2.3). These species typically have distinct spectral features (Figure 2.S2) that allow for strong constraints. The retrieved LOS abundances of these species are all within ~1-σ of the true values, indicating that our retrieval method is stable to random noise. The LOS abundance PDFs of some other species (e.g., C_2H_6, C_6N_2, haze and C_2N_2 in Figure 2.3) show asymmetric behavior, and so can only

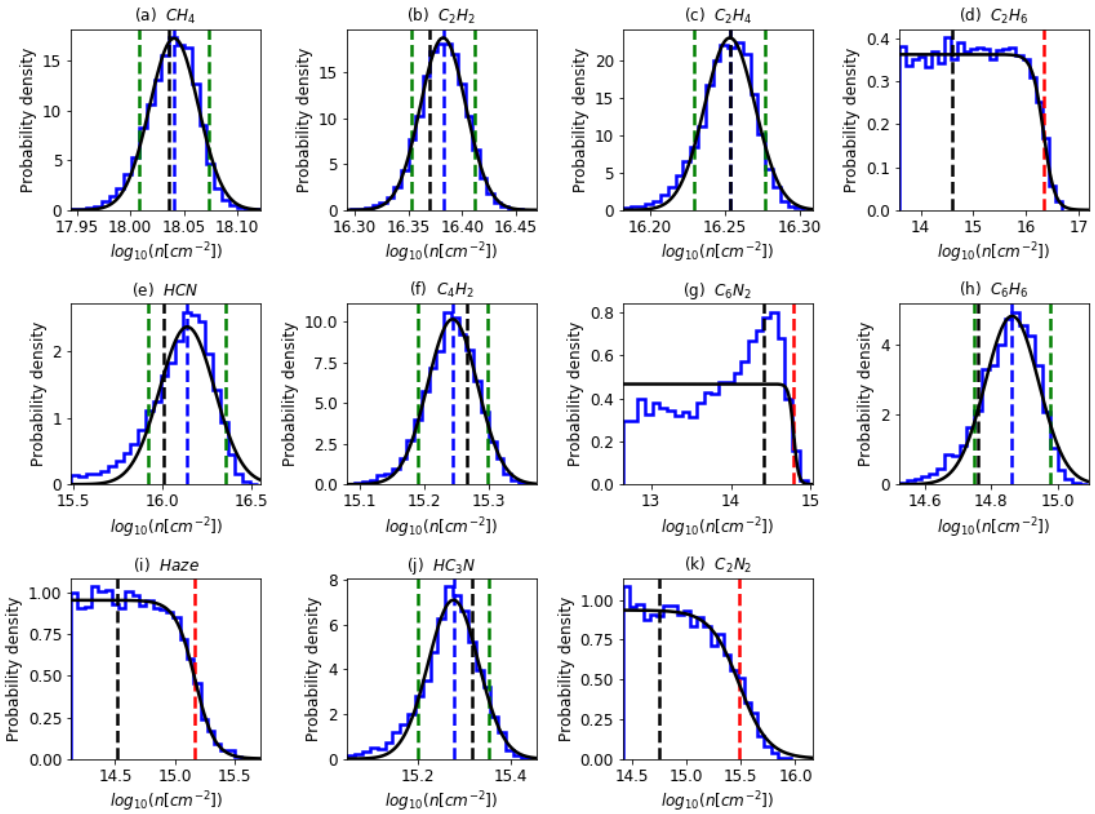

Figure 2.3. Probability density functions (PDFs, blue solid lines) of the logarithm of LOS abundances retrieved from the synthetic spectrum (red line in Figure 2.2). The black dashed lines indicate the LOS abundances used to generate the synthetic spectrum, i.e. the "true" value. The best fit function to each PDF (Gaussian, sigmoid, and constant) is shown as black solid lines. The median and 1-σ confidence interval are denoted by blue and green dashed lines, respectively, for well-constrained species. For others, the upper limits are denoted by red dashed lines. The correlations among the parameters are shown in Figure 2.S3 in the appendix.

provide upper limits through fits to the sigmoid function. Interestingly, ethane (C_2H_6), one of the major hydrocarbons, belongs in this category due to overlapping spectral features with the most abundant hydrocarbon, methane (CH_4, Figure 2.1d), whose LOS abundance is two orders of magnitude higher, resulting in the anti-correlation of these two PDFs (Figure 2.S3). The failure to retrieve ethane is consistent with previous results obtained above 700 km during flyby T41i (Koskinen et al. 2011). In some cases, the PDFs may be between Gaussian and sigmoid, such as that of C_6N_2 in this analysis, which shows a peak with poorly constrained lower limits; as the type of PDF is still determined by the value of

the residuals, the case of C_6N_2 and others like it would either have only upper limits (more sigmoid-like) or be constrained with relatively large uncertainties (more Gaussian-like).

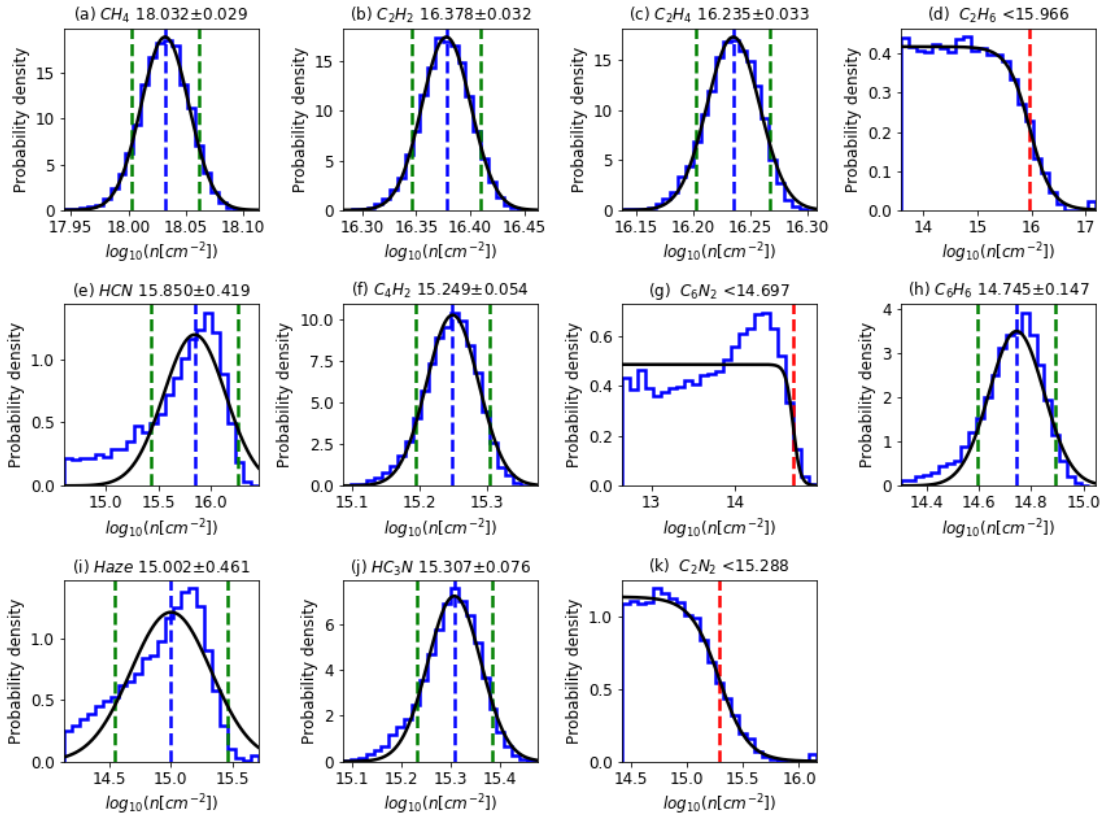

Figure 2.4. Probability density functions (PDFs) of the logarithm of LOS abundances retrieved from a photon count spectrum at 754 km ray tangent height (blue solid lines). The best fit function to each PDF (Gaussian, sigmoid, and constant) is shown as black solid lines. The median and 1-σ confidence interval are denoted by blue and green dashed lines, respectively, for well-constrained species. For others, the upper limits are denoted by red dashed lines. The medians, 1-σ confidence intervals, and upper limits are also given atop each subplot. The correlations among the parameters are shown in Figure 2.S4 in the appendix.

2.6 Results and discussion

An example spectrum from the T52 occultation observation at ~750 km is shown in Figure 2.2, together with a simulated spectrum that best fits the observation. PDFs for the LOS

abundances of hydrocarbon and nitrile species resulting from retrievals of this spectrum are shown in Figure 2.4. Most of the behavior of these PDFs are consistent with those in Figure 2.3. Six species (CH_4, C_2H_2, C_2H_4, C_4H_2, C_6H_6, and HC_3N) show Gaussian-like PDFs, and are thus well constrained with precise values for the LOS abundances and small uncertainties. It is worth noting that the uncertainties shown here are only from photon noise; another probable source of uncertainties is the presently unavailable temperature dependencies of extinction cross sections, which may introduce systematic errors. Three other species (C_2H_6, C_6N_2, and C_2N_2) have asymmetric PDFs more similar to the sigmoid function, with only well-defined upper limits. HCN and haze both have PDFs that are close to Gaussian, and as such are categorized as well constrained, but with large uncertainties (about a factor of 3). The third PDF type, a constant, is not shown in either Figure 2.3 or Figure 2.4, as all retrieved species are constrained to some extent at ~750 km. A constant PDF usually takes place only when the ray tangent height is either too high or too low in the atmosphere, where the LOS abundances of some species are either insufficient to be seen or overwhelmed by saturated absorption, respectively. In other words, flat PDFs would be returned when there is almost no information in the spectrum.

Repeating the procedures outlined above for all available altitudes, we present the vertical profiles of LOS abundances of all eleven species in Figure 2.5 with error bars and arrows for well-defined constraints and upper limits, respectively. This is the first time that the LOS abundance profiles of these species are retrieved from a flyby with significant pointing motion. Comparison of the CH_4, C_2H_2, C_2H_4, and HCN profiles with those retrieved from the TB (Shemansky et al. 2005) and T41i flybys (Koskenin et al. 2011) are given in Figure 2.S5. It shows general agreement between these three retrievals despite some differences at higher altitudes, which may result from the different seasons and/or latitudes. Among the major hydrocarbons (Figure 2.5a), only the most abundant species, CH_4, is detectable above 1000 km, while all the others only exhibit upper limits. With decreasing altitude, concentrations of larger organic molecules increase due to photochemistry. All of the major hydrocarbons except C_2H_6 remain well constrained down to 400 km, where absorption saturates. Spectral features of C_2H_6 overlap with that of CH_4

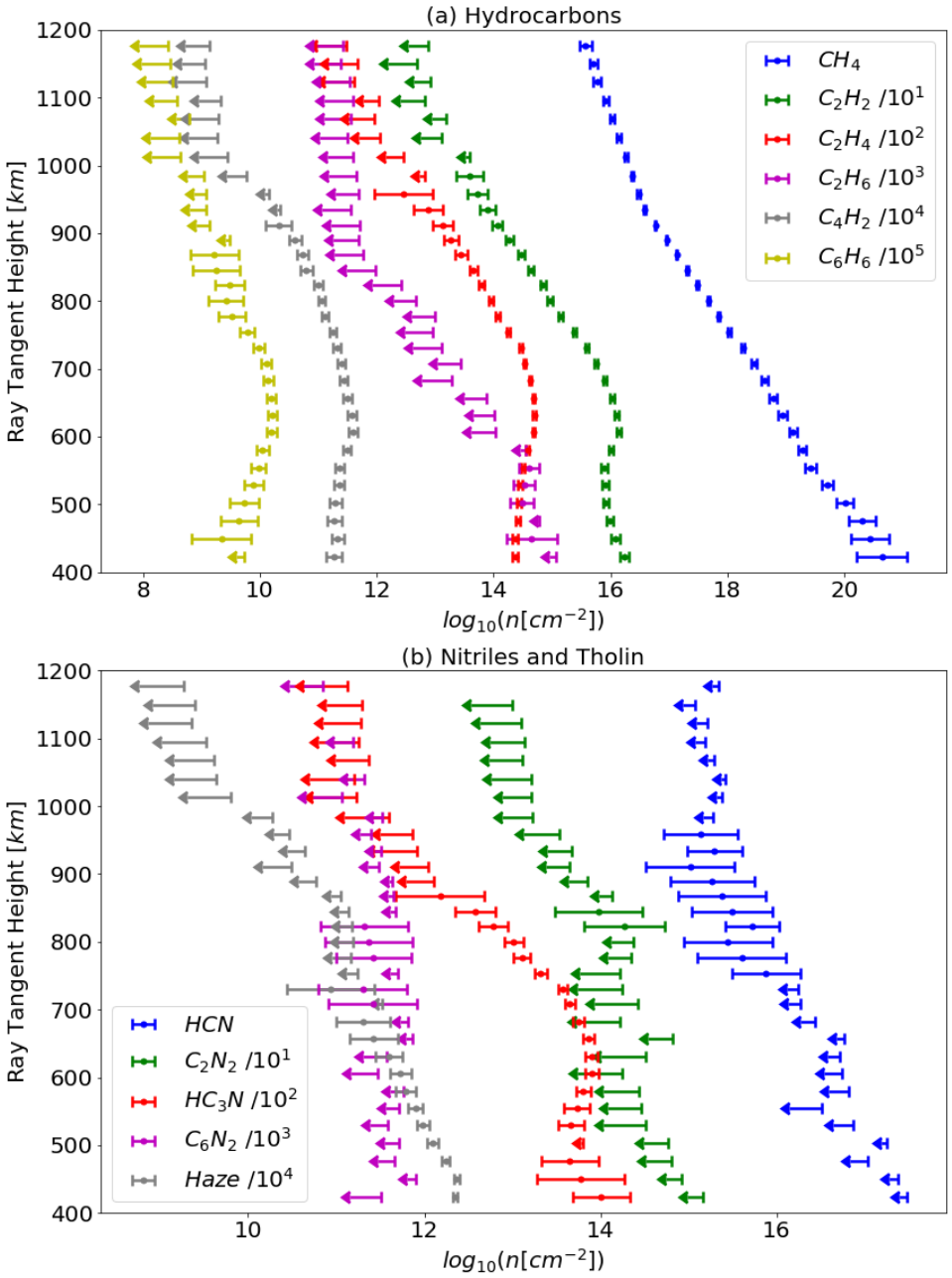

Figure 2.5. Vertical profiles of the logarithm of LOS abundances retrieved from T52 occultation observations. Some species are offset by a few orders of magnitude for the purpose of presentation. Points with error bars denote well-constrained values, while arrows denote upper limits. The lengths of the arrows denote the width of each soft upper limit threshold. Haze particles are assumed to be 12.5 nm spheres with the same optical properties as their laboratory analog ("tholin"; Khare et al. 1984). Comparison of LOS abundances of CH_4, C_2H_2, C_2H_4 and HCN with two previous flybys (TB and T41i) is given in Figure 2.S5.

for most of the altitude range, resulting in a failure to retrieve its abundance, as discussed in Section 2.4. Among nitriles, HCN is constrained with relatively large uncertainties between 700 and 1000 km, while HC_3N is well constrained to much lower altitudes. Only upper limits for C_6N_2 and C_2N_2 are obtained, and so we do not claim a detection. Tighter constraints may be obtained from spectra with higher signal to noise ratio measured during a more stable flyby (e.g., T41i), which we will investigate in a future publication. Aside from the major hydrocarbon and nitriles, our new method can also identify the two long wavelength absorbers, benzene and haze. Benzene is well constrained below 900 km due to its distinguishable feature near 1790 Å, while haze particles can only be retrieved below 750 km. The range of altitudes where haze is well constrained in our work is smaller than those of Liang et al. (2007) and Koskinen et al. (2011). Liang et al. (2007) retrieved the haze profile up to 1000 km from observations obtained during flyby TB by assuming that all extinction between 1850 to 1900 Å is caused by haze, as benzene was not included as a potential absorber. However, the current work shows that benzene can contribute up to 50% of the extinction in this wavelength range, so the ambiguity of the two absorbers needs to be considered. Koskinen et al. (2011) also retrieved the haze profile up to 1000 km using observations obtained during two stable flybys T41i and T53, while assuming that the PDF of the LOS abundance of haze was Gaussian. Both the higher signal to noise ratio of these observations and the assumption of a Gaussian PDF could have contributed to a greater altitude range where haze LOS abundance can be constrained. Applying our method to these flybys may help to understand why these differences exist.

LOS abundances of eight species (CH_4, C_2H_2, C_2H_4, HCN, C_4H_2, C_6H_6, haze, and HC_3N) are converted to number density profiles to allow for ease of comparison to photochemical models. Three species (C_2H_6, C_2N_2 and C_6N_2) are excluded, as their LOS abundances are not well constrained over most of the considered altitude range. The Abel inverse transform is used here, which assumes spherical symmetry, to compute the vertical profiles. A Bootstrap Monte Carlo (BSMC) method is used to evaluate the quality of conversion and provide uncertainties. BSMC has the advantage of being applicable to different types of

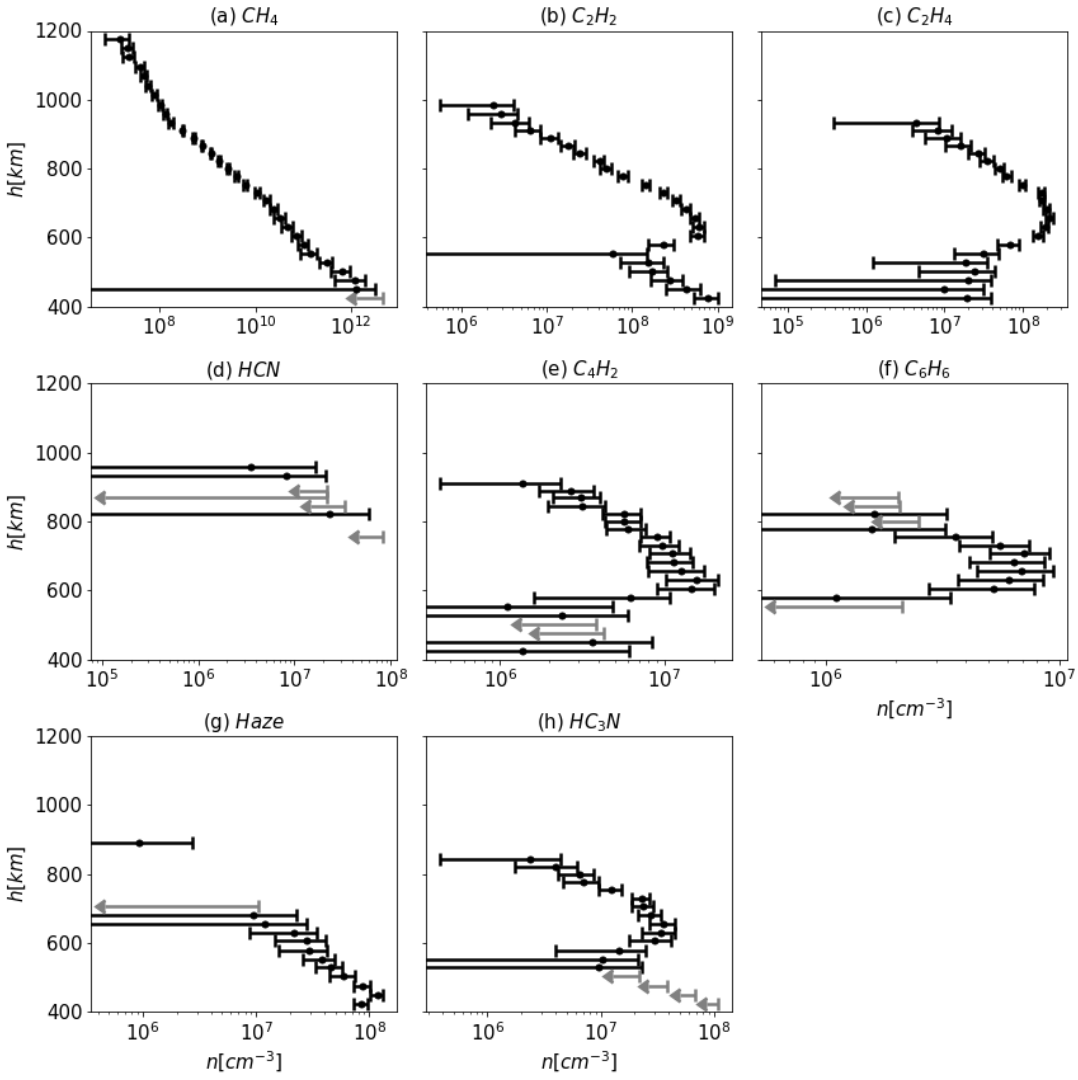

Figure 2.6. Number density profiles of selected species retrieved from T52 occultation observations. Black points with error bars denote well-constrained values, while grey arrows denote upper limits. The lengths of the arrows denote the width of each upper limit soft threshold. Haze particles are assumed to be 12.5 nm spheres with the same optical properties as their laboratory analog ("tholin"; Khare et al. 1984).

PDFs of the LOS abundances, which is necessary since a number of the PDFs are not Gaussian. The number density profiles of each species are computed individually since each species is independent of others. In each computation, a set of LOS abundances at all retrieving altitudes for the given species is sampled from their PDFs (e.g. one of the panels in Figure 2.4) at each BSMC step, from which we compute the corresponding vertical

number density profile. As the distribution of species LOS abundances in the Markov chains obtained from the retrieval are identical to their PDFs when the MCMC procedure reaches equilibrium, we used the values in the last 1500 steps of each of the 120 chains as the sampling procedure for BSMC and generate 180000 number density profiles for each species. Therefore, at each altitude of an individual species, we obtain 180000 probable number densities, which form a number density PDF. To interpret these PDFs, we use the same method of fitting them with three types of functions (Gaussian, sigmoid, and constant) and categorized them by comparing the residuals as mentioned in Section 2.4. Number density profiles corresponding to well-constrained number densities and upper limits with positive values are shown in Figure 2.6. As the number density at each altitude is computed with contributions from both well and poorly constrained LOS abundances, the number densities have larger relative uncertainties and smaller ranges of altitudes where they are well constrained. Major hydrocarbons with large abundances and distinct spectral features (CH_4, C_2H_2 and C_2H_4; Figure 2.6a-2.6c) are well constrained over a wide range of altitudes. In contrast, the number density of the most abundant nitrile, HCN, is poorly constrained for most of the altitude range considered (Figure 2.6d) due to its large LOS abundance uncertainty (Figure 2.5). Other minor species (Figure 2.6e-2.6h) are well constrained over at least some of the considered altitude range, and can thus provide constraints on Titan's atmospheric chemistry.

2.7 Conclusions

A new method to correct for the effects of pointing motion of Cassini/UVIS has been developed using an instrument simulator and the MCMC method. The new approach is successfully applied to the T52 stellar occultation observations of Titan's atmosphere to retrieve the LOS abundances and number densities of hydrocarbon and nitrile species, and allows for the quantification of how well each LOS abundance and number density can be constrained, facilitating the analysis of all Cassini/UVIS stellar occultations at Titan. Application of the present method to all available observations is expected to reveal seasonal and latitudinal variations in the atmospheric composition of Titan, thereby

providing useful constraints for photochemical and global circulation models.

2.8 Appendix

The influence of pointing motion on the distortion of the photon count spectrum of T52 at 754 km ray tangent height is shown as Figure 2.S1. It is the same as Figure 2.1c in the main text, but expanded to include the entire FUV wavelength range.

The spectral extinction contribution of each species to the spectrum of T52 at 754 km ray tangent height is shown as Figure 2.S2. Even though the zero-extinction spectrum does not exhibit a pointing drift, we add that of the spectrum at 754 km to allow for better inter-comparison.

Correlations among the probability density functions (PDFs) shown in Figures 2.3 and 2.4 in the main text are plotted as Figures 2.S3 and 2.S4, respectively. Numbers and units in Figures 2.S3 and 2.S4 are the same as those shown in Figures 2.3 and 2.4.

The LOS abundances of CH_4, C_2H_2, C_2H_4, and HCN retrieved from T52 occultation observations and shown in Figure 2.5 are compared with those from two previous flybys (TB; Shemansky et al. 2005 and T41i; Koskinen et al. 2011) in Figure 2.S5.

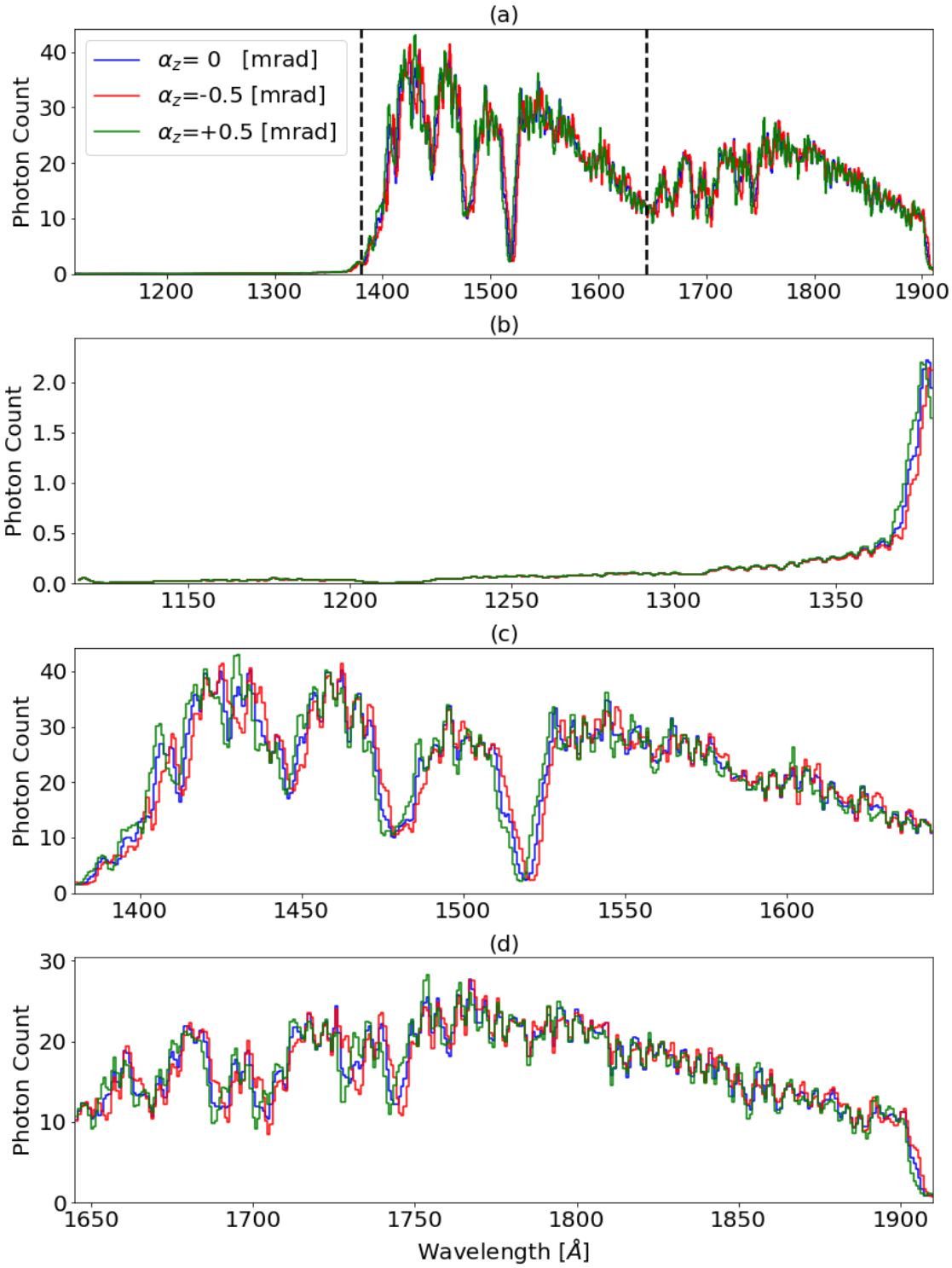

Figure 2.S1. (a) Simulated photon count spectra from the T52 occultation at 754 km ray tangent height with pointing motion values along the spectral dimension of 0 (blue), -0.5 (red) and +0.5 (green) mrad. (b)-(d) Detailed views of the photon count spectra split along the black dashed lines in (a). The y-axes scales are different in (b)-(d).

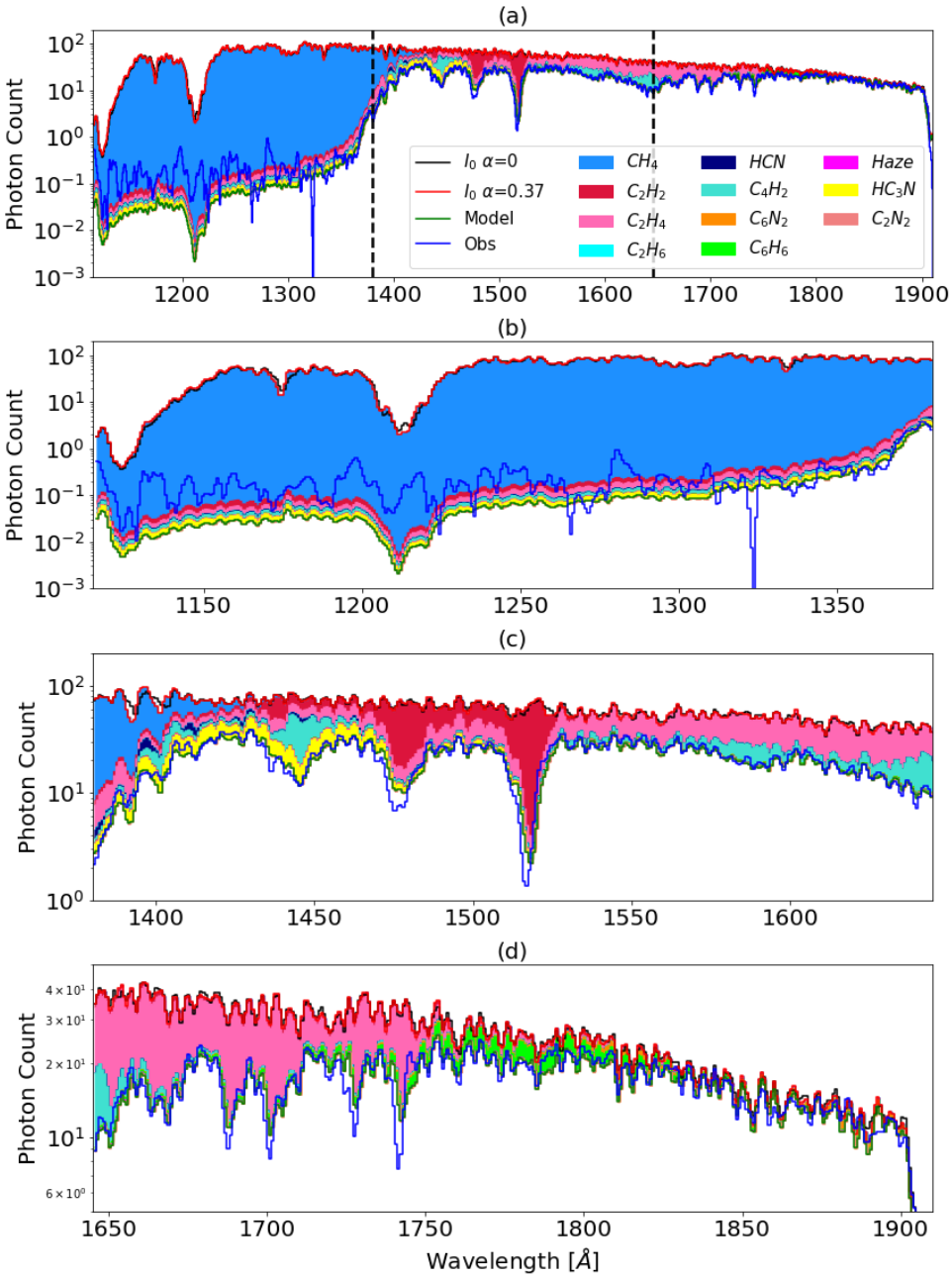

Figure 2.S2. (a) Photon count spectra from the T52 occultation. The spectra observed above Titan's atmosphere (I_0) without and with a pointing drift of 0.37 mrad, its value at 754 km ray tangent height, are shown as black and red solid lines, respectively. The observed and best-fit model spectra at 754 km ray tangent height are shown as blue and green solid lines, respectively. The extinction contributions of each species are shown by the shaded areas between spectra. The y-axis is in log scale such that the area denoted for each species is proportional to their contributions to optical depth. (b)-(d) Detailed views of the photon count spectra split along the black dashed lines in (a). The y-axes scales are different in (b)-(d).

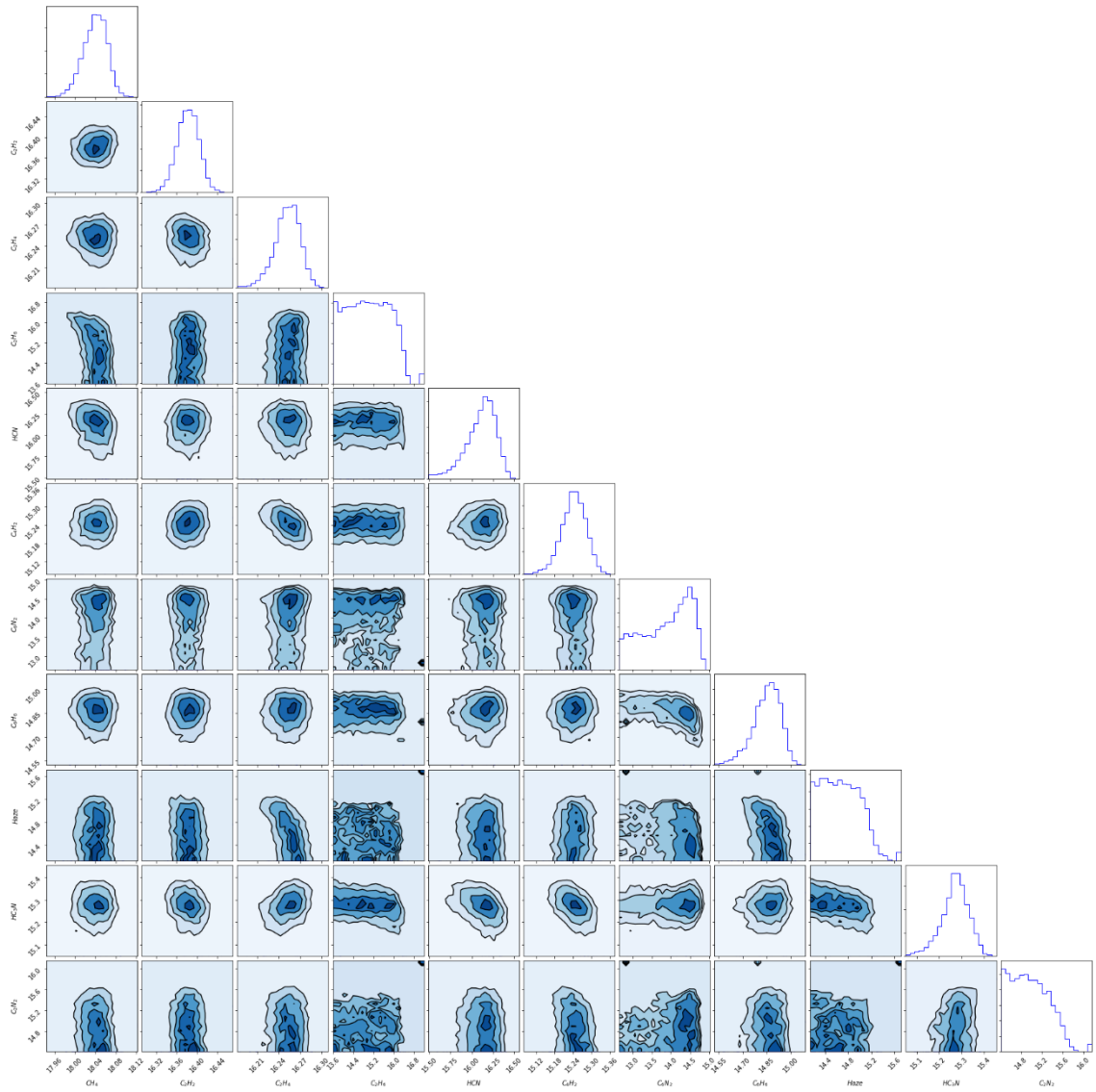

Figure 2.S3. Probability density functions (PDFs) of the logarithm of LOS abundances, which are the same as those shown in Figure 2.3, and the correlations among them retrieved from the synthetic spectrum shown in Figure 2.2. Axes units are the same as those in Figure 2.3.

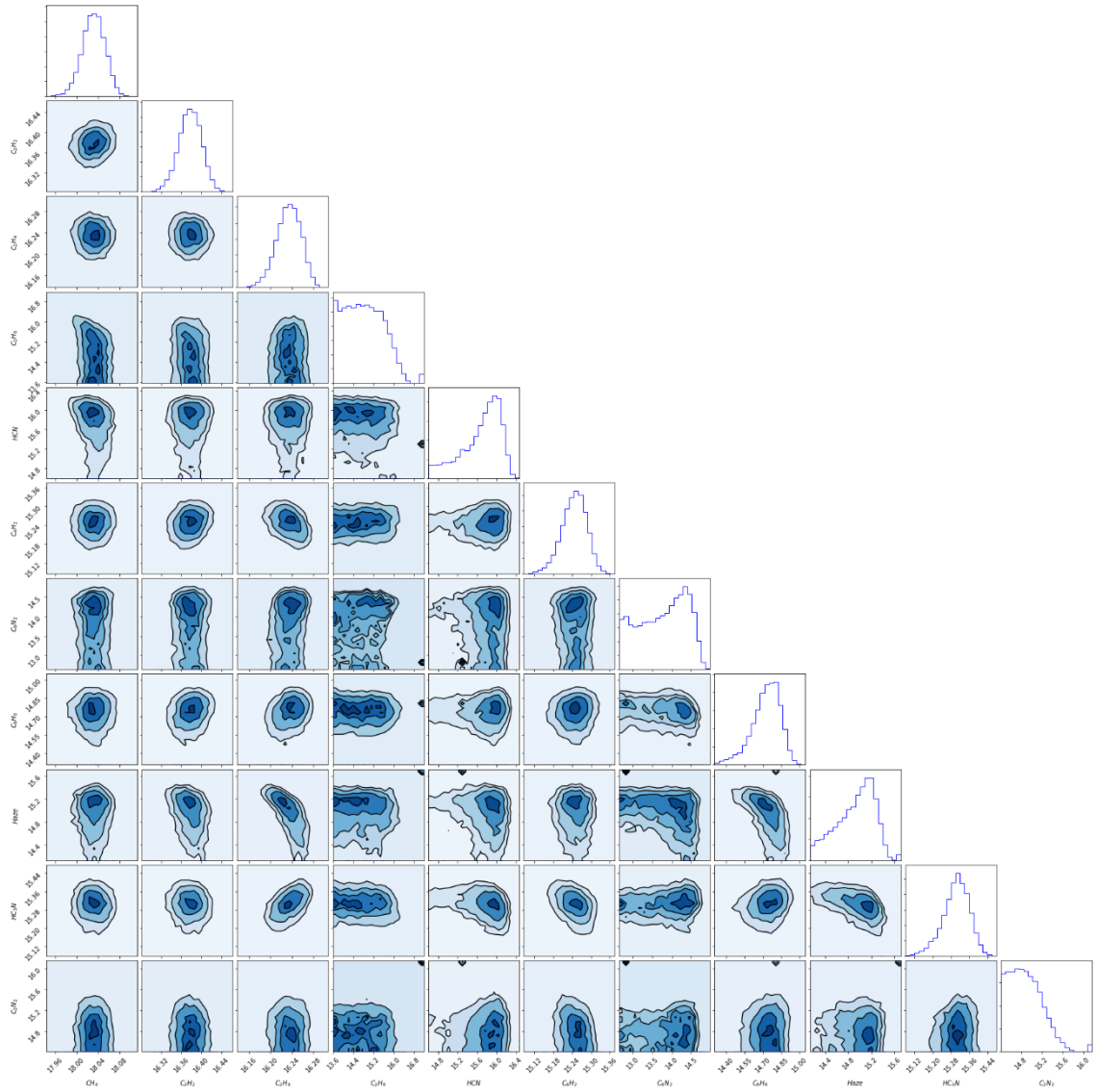

Figure 2.S4. Probability density functions (PDFs) of the logarithm of LOS abundances, which are the same as those shown in Figure 2.4, and the correlations among them retrieved from the observed photon count spectrum at 754 km ray tangent height shown in Figure 2.2. Axes units are the same as those in Figure 2.4.

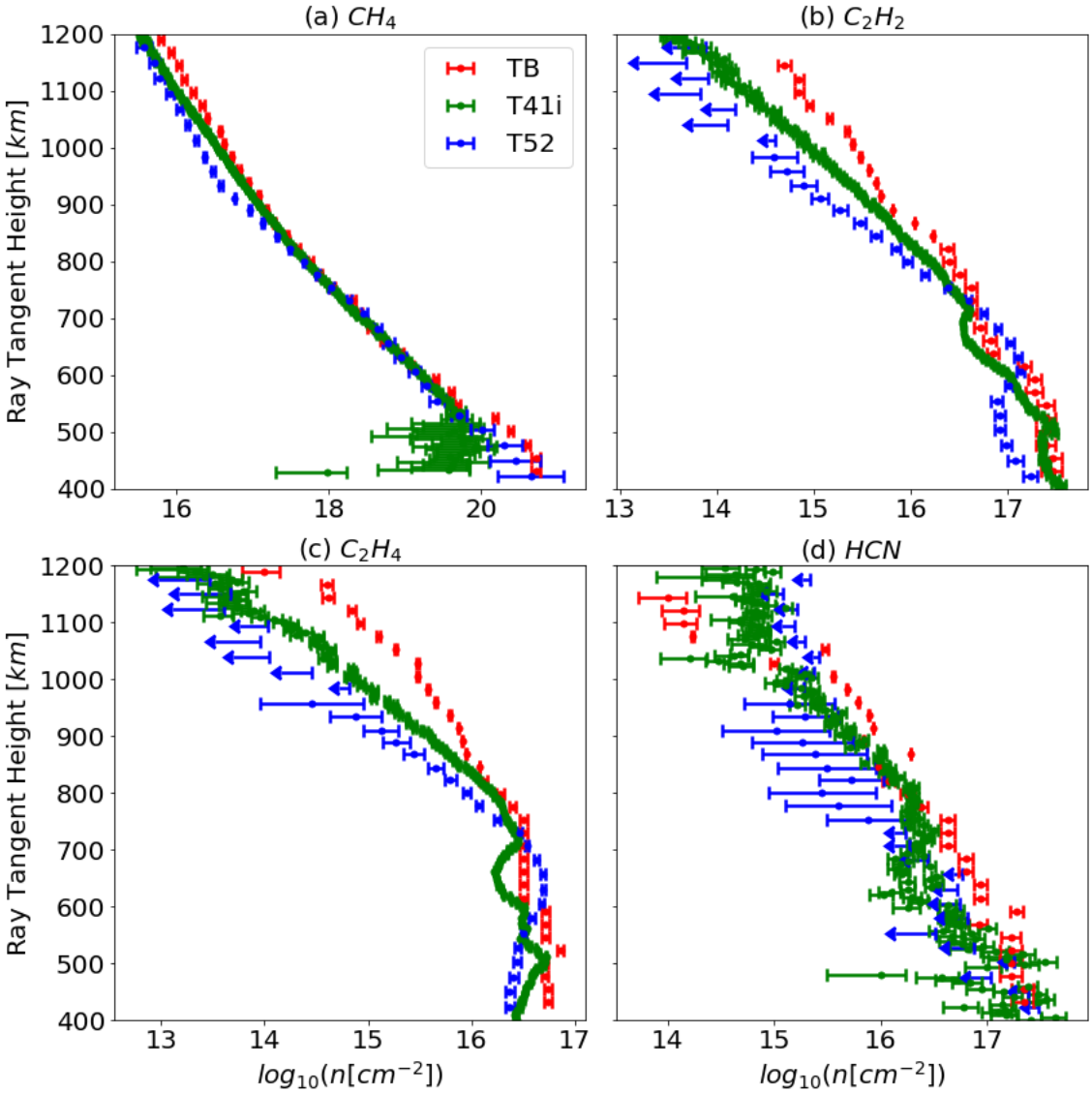

Figure 2.S5. Vertical profiles of the logarithm of LOS abundances retrieved from T52 occultation observations (blue), compared to those retrieved from observations during TB (red; Shemansky et al. 2015) and T41i (green; Koskinen et al. 2011) by the cited works.

Chapter 3

A bimodal distribution of haze in Pluto's atmosphere

Siteng Fan[1], Peter Gao[2], Xi Zhang[2], Danica J. Adams[1],

Nicholas W. Kutsop[3], Carver J. Bierson[4], Chao Liu[1],

Leslie A. Young[5], Andrew F. Cheng[6], Yuk L. Yung[1,7]

[1]California Institute of Technology, Pasadena, CA 91125.

[2]University of California, Santa Cruz, CA 95064.

[3]Cornell University, Ithaca, NY 14850.

[4]Arizona State University, Tempe, AZ 85287.

[5]Southwest Research Institute, Boulder, CO 80302.

[6]Applied Physics Laboratory, Johns Hopkins University, Laurel, MD 20723.

[7]Jet Propulsion Laboratory, California Institute of Technology, Pasadena, CA 91109.

3.1 Abstract

We present a joint retrieval of Pluto's haze morphology and distribution in the lower 50km of its atmosphere, using observations from multiple instruments onboard the New Horizons spacecraft. Haze in Pluto's atmosphere is one of the most distinctive features seen during New Horizons' historic flyby in 2015. The haze was directly imaged at visible and near infrared wavelengths by the Long Range Reconnaissance Imager (LORRI), the Multispectral Visible Imaging Camera (MVIC), and the Linear Etalon Imaging Spectral Array (LEISA), and investigated in the ultraviolet (UV) by the Alice spectrograph using solar occultations. Contradictory conclusions have been drawn from observations by the different instruments. For example, forward and backward scattering at visible wavelengths cannot be reconciled with the observed UV extinction if only one haze particle population were assumed. Here, we present retrieval results of a bimodal distribution of haze particles, which consists of a small population with radius ~80nm and a large population with radius ~1μm. This result successfully addresses the disagreement among the instruments, and provides important constraints on local haze production and gas condensation in Pluto's atmosphere.

3.2 Introduction

The existence of haze in Pluto's atmosphere was confirmed by the New Horizons spacecraft during its flyby in July 2015 (Stern et al. 2015, Gladstone et al. 2016). Hypothesized to originate from photolysis of methane, nitrogen, and larger organic molecules in Pluto's upper atmosphere, haze particles are thought to grow through coagulation of smaller particles as they sediment downwards (Gao et al. 2017), similar to processes in Titan's atmosphere (Lavvas et al. 2013). Haze plays an important role in the Pluto system. The macromolecular particles serve as the condensation and/or sticking surface for chemical species in Pluto's atmosphere (Wong et al. 2017; Luspay-Kuti et al. 2017). They may also dominate Pluto's atmospheric energy budget in two ways. Haze particles likely significantly influence the temperature through radiative heating and cooling, which explains Pluto's unexpected cold atmosphere (Zhang et al. 2017). After condensing onto the surface, haze particles alter its composition and therefore albedo by becoming darker due to solar radiation (Grundy et al.

2018).

Despite the important role of haze in the Pluto system, however, the morphology of haze particles still remains a mystery to date. Observations of haze particles obtained by the Multispectral Visible Imaging Camera (MVIC, Reuter et al. 2008) at optical wavelengths present a bluish color, which suggests small particle radii compared to the observation wavelength (Gladstone et al. 2016). On the other hand, observations also show a strong forward scattering by haze (Gladstone et al. 2016, Cheng et al. 2017), which is an indication of large particles. Given these properties, the haze particles are likely fractal aggregates, highly porous and randomly shaped ~μm particles consisting of small ~10nm spheres, similar to haze in Titan's atmosphere (Khare et al. 1984). Gao et al. (2017) used a numerical microphysics model to investigate haze formation and interpret the ultraviolet (UV) extinction vertical profiles obtained by the Alice UV Imaging Spectrograph (Stern et al. 2008, Young et al. 2018). However, disagreement exists between the UV extinction and scattering phase functions at visible wavelengths obtained by the LOng-Range Reconnaissance Imager (LORRI, Cheng et al. 2008), such that no solution has been found that can explain both of sets of observations (Cheng et al. 2017). Besides observations by these three instruments, infrared (IR) spectra of haze particles at a forward scattering configuration were obtained by the Linear Etalon Imaging Spectral Array (LEISA, Reuter et al. 2008, Grundy et al. 2018), which have not been evaluated. Therefore, there is no unique solution for the morphology of Pluto's haze that has been found to date that can explain the available observations. As a result, the nature of Pluto's haze, and its importance to the Pluto system are loosely constrained. This is the subject of this work.

In this work, we present a joint retrieval using observations obtained by all instruments onboard New Horizons spacecraft that observed Pluto's haze, and conclude that a bimodal distribution of haze is the only solution after exhausting a number of scenarios. In Section 3.3, we summarize the available observations of Pluto's haze used in this work. The methodology including details of the forward model and parameter searching approach in the retrieval is shown in Section 3.4. A number of scenarios of haze particles morphology

are evaluated in Section 3.5. Discussion and summary follow as Section 3.6 and 3.7.

3.3 Observations and Data Processing

Observations of Pluto's haze were obtained by multiple instruments onboard the New Horizons spacecraft during the flyby with wavelengths ranging from UV to IR. A summary of observations used in this work is given in Table 3.1 and Figure 3.1.

Table 3.1. Summary of Pluto's haze observations.

Instrument	Wavelength (μm)	Altitude Range (km)	Phase Angle (degree)	Reference
Alice	0.165	0-300	Extinction	Young et al. (2018)
LORRI	0.608	0-100	19.5	Cheng et al. (2017)
		0-50	67.3	
		0-75	148.3	
		0-200	169.0	
LEISA	1.235-2.435	0-299	169.0	Grundy et al. (2018)
MVIC	0.624 0.492 0.861 0.883	0-50	18.2 38.8 169.4	This work

3.3.1 Alice

The Alice UV spectrograph is an imaging spectrograph with a bandpass from 52 to 187nm excluding a gap near Ly-α from 118 to 125nm (Stern et al. 2008). It obtained two UV extinction vertical profiles of Pluto's atmosphere from ingress and egress of a solar occultation (Young et al. 2018). These two profiles are consistent despite the ingress (15.5 °S, 195.3°E) and egress (16.5°N, 13.3°E) points being on opposite sides of Pluto's limb. In this work, we adapt the two haze extinction profiles that are derived and analyzed in detail in Young et al. (2018). 185nm is selected to be the representative wavelength as the haze's extinction cross section varies less than 3% at wavelengths longer than 180nm when the haze's optical constants are assumed to be the same as those of Titan tholins (Khare et al. 1984). The extinction vertical profiles are binned every 5km, and uncertainties are computed through error propagation (Figure 3.1a).

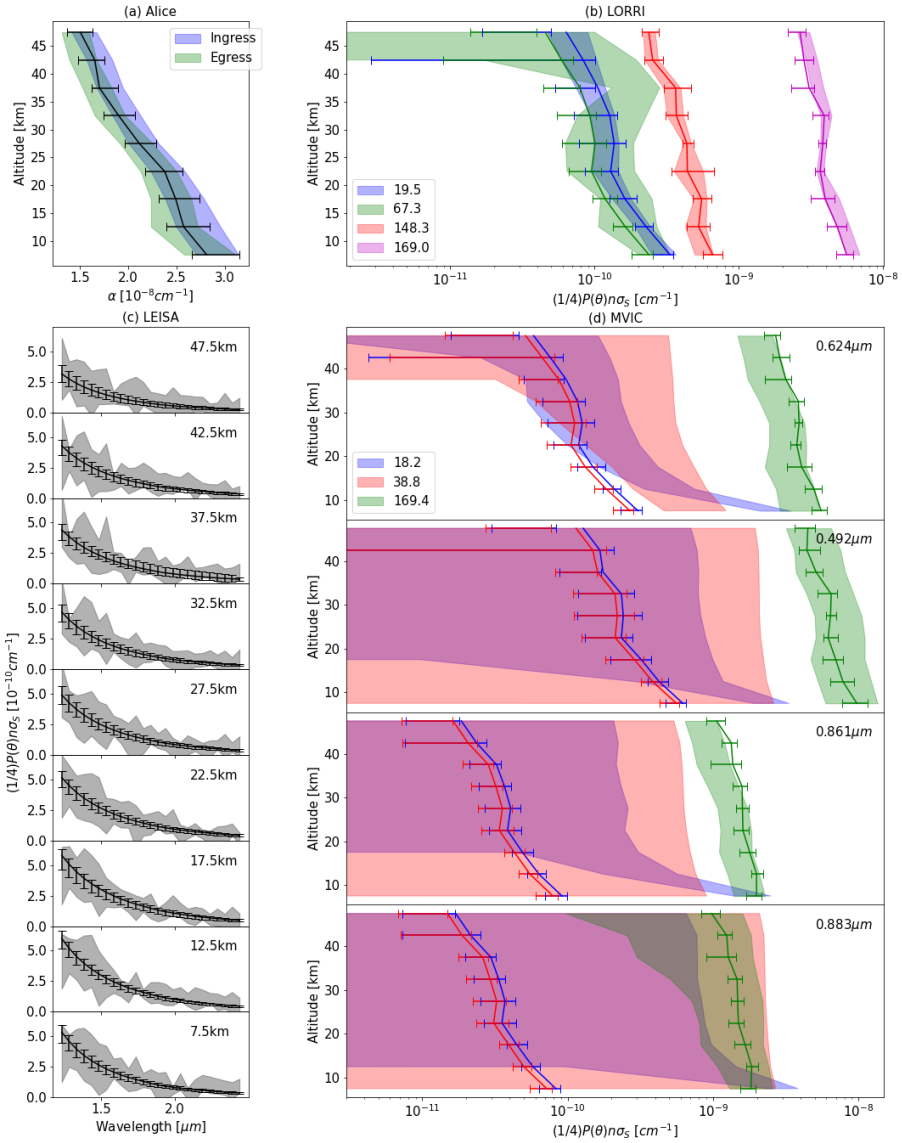

Figure 3.1. Comparison between observations obtained by instruments onboard New Horizons spacecraft (shaded areas) and model results with 1-σ uncertainties under the scenario of a bimodal distribution of haze particles (solid lines with error bars). (a) UV extinction coefficient at 0.185μm obtained by Alice spectrograph during solar occultation ingress (blue) and egress (green), taken from Young et al. (2018). (b) Local scattering intensity at 0.608μm derived from LORRI images at four phase angles of 19.5 °(blue), 67.3 °(green), 148.3 °(red), and 169.0 °(magenta), processed using data from Cheng et al. (2017). (c) Local scattering intensity spectrum of each altitude bin at 1.235-2.435μm derived from LEISA observations at phase angle of 169.0 °, processed using data from Grundy et al. (2018). (d) Local scattering intensity at 0.624, 0.492, 0.861, 0.883μm (from top to bottom) wavelength bands derived using MVIC images at three phase angles of 18.2 °(blue), 38.8 °(red), 169.4 °(green).

3.3.2 LORRI

LORRI is a panchromatic visible imager with a wide bandpass from 350 to 850nm, and a pivot wavelength of 607.6 nm (Cheng et al. 2008). It imaged Pluto's haze at a number of different phase angles. In these images, haze was found to be global and forming thin layers, and no temporal change was identified (Cheng et al. 2017). I/F profiles in images that are taken at several mean phase angles from 19.5 ° to 169.0 ° were extracted by Cheng et al. (2017) to analyze the property of haze particles. However, no solution was found that can simultaneously match both these I/F profiles and the UV extinction obtained by Alice. Here, we adapt four I/F profiles from Cheng et al. (2017) with the highest resolution into this work, covering phase angles from backward to forward scattering. The I/F profiles are binned to altitude bins with the same width as Alice, and uncertainties are computed as the standard deviation of I/F in each bin. As the I/F in the images are line-of-sight (LOS) integrated quantities, we use the Abel transform for noisy data to derive local densities (Figure 3.1b), whose details are given in Appendix 3.8.1. Given the small optical thickness of Pluto's atmosphere, we assume the single scattering approximation and spherical asymmetry in the transform, although there exist some variations over large spatial scales (Cheng et al. 2017). As I/F values above the available observed altitudes are required in the transform, we fit a decreasing exponential function in geopotential, in line with Young et al. (2018).

3.3.3 LEISA

LEISA is a short-wavelength IR spectral imager in the Ralph instrument (Reuter et al. 2008). Its primary filter covers a wavelength range from 1.25 to 2.5μm. Despite its primary purpose of mapping the surface geology and composition, LEISA obtained an observation of Pluto's haze at a forward scattering geometry with a phase angle of 169 °(Grundy et al. 2018). We adapt the I/F profiles and spectra extracted from this image to place IR spectral constraints on the haze particles. The LOS integrated I/F profiles are also transformed to local scattering densities using the Abel transform (Figure 3.1c), and corresponding uncertainties are computed through error propagation (Appendix 3.8.1).

3.3.4 MVIC

MVIC is a visible, near-infrared (NIR) imager in the Ralph instrument (Reuter et al. 2008). Besides optical navigation, it takes color images of the Pluto system using color arrays under time delay integration (TDI) mode. This allows it to obtain images in the blue (400-550nm), red (540-700nm), NIR (780-975nm) and a narrow methane absorption band (860-910nm). During the flyby of New Horizons, MVIC took a number of images of Pluto's haze under this color TDI mode at different phase angles, which constitute a group of constraints. We select three observations with the highest resolutions, which are better than 5 km/pixel at mean phase angles of 18.2 ° 38.8 ° and 169.4 °. The mission elapsed time (MET) of these observations are 0299162512, 0299178092, and 0299193157. Details of the MVIC data processing are given in Appendix 3.8.2, including the correction of Pluto's disk glow and the extraction of the I/F profiles from the images. We compute the vertical I/F profiles and their uncertainties using the Abel transform (Figure 3.1d, Appendix 3.8.1), the same as that for the LORRI and LEISA observations. Kutsop et al. (in review) investigated haze properties using observations at three of the color filters and at seven phase angles across a much greater altitude range but at lower altitude resolution; they suggest that Pluto's haze particles follow either a bimodal distribution or a power-law distribution, and that a monodisperse population may not be feasible. We combine these observations with those from other instruments to address the degeneracy between these two distributions.

Figure 3.2. Illustration of fractal aggregate morphology with fractal dimensions (D_f) of 1.8, 2.0, and 2.2 (from left to right). The number of monomers is 100 in each of the aggregate.

3.4 Methodology

Through data processing, we obtain extinction in the UV, scattering intensities at a few phase angles in the visible and NIR, and a scattering intensity spectrum in the IR at each altitude bin. The bins range from the surface to 50 km with a width of 5 km. To derive the haze morphology from these observations, we solve a multivariable inverse problem with a combination of a forward model and a retrieval algorithm.

3.4.1 Aggregate Morphology

Given the bluish color and strong forward scattering, haze particles are suggested to be fractal aggregates, highly porous and randomly shaped particles consisting of small subunits, "monomers". To parameterize the morphology, we assume that the monomers within an aggregate are identical spheres, and the fractal aggregate can be described by three parameters, the fractal dimension (D_f), monomer radius (r_m) and number of monomers in each aggregate (N_m). The fractal dimension controls the porosity of the aggregate and relates the change in size of the aggregate with its change in mass. Figure 3.2 visually shows the influence of D_f on aggregate morphology. Aggregates with larger D_f are more compact. The relationship between aggregate effective radius (R_a) and these three parameters is

$$N_m = \left(\frac{R_a}{r_m}\right)^{Df} \qquad (3.1)$$

3.4.2 Scattering Models

Light scattering of spherical haze particles and monomers in fractal aggregates is computed using Mie theory (Wiscombe 1979), as their radii are comparable to some of the observation wavelengths. The complex refractive index of these spheres, which is determined by the particle chemical composition, are assumed to be the same as those in Titan's atmosphere due to the similarity between these two atmospheres, which are both nitrogen dominant atmospheres with trace organic molecules (Gladstone et al. 2016). We adapt the measurement of Khare et al. (1984) where optical properties of tholin-like particles were

produced under a pressure of 0.2mbar in the laboratory. Another more recent laboratory measurement (Ramirez et al. 2002) is not used in this work due to its high haze particle production pressure (2mbar), two orders of magnitude higher than Pluto's surface pressure (~0.01mbar, Gladstone et al. 2016).

We use the scattering model described by Tomasko et al. (2008) to estimate the phase function and cross sections of the parameterized fractal aggregate particles. The model was originally developed to constrain haze particle properties in Titan's atmosphere. The scattering model uses empirical phase functions derived from averaging exact results of randomly produced aggregates, which significantly reduces the computational time, resulting in the feasibility of retrieval. It computes the phase function and cross sections at a given wavelength with three parameters describing an aggregate (D_f, r_m, and N_m) and the complex refractive index as described above. Another scattering model (Rannou et al. 1999) was tested but not used in this work due to its omission of polarization, which is not negligible when monomer number becomes large (~1000), although its disagreement with Tomasko et al. (2008) is less than 20% in our interested parameter space. The scattering model we used in this work is rigorously tested at D_f=2 and mostly at $N_m<10^3$, but testing has shown that perturbation of D_f is allowed ($1.5<D_f<2.5$), and the extrapolation of N_m to ~10^4 is reasonable given the linear relationship in log-log scale between cross section and monomer number at the larger end of aggregate size (Tomasko et al. 2008).

3.4.3 Retrieving Algorithm

We choose the Markov chain Monte Carlo (MCMC) method (e.g. Foreman-Mackey et al. 2013) as the parameter searching tool. It derives the posterior probability density function (PDF) of each parameter by comparing the posterior probabilities of proposed parameter sets. The cost function with posterior probability (p) is defined as follows which includes the value (μ_i) and uncertainty (σ_i) of each observation.

$$\ln(p) = -\frac{1}{2}\sum_i \frac{(X_i-\mu_i)^2}{\sigma_i^2} \qquad (3.2)$$

where i is the index of the i-th observation, and X_i is the modeled i-th observation computed using a given proposed parameter set during one MCMC attempt. The Python package *emcee* (Foreman-Mackey et al. 2013) is used to implement the parameter searching algorithm. We initiate the MCMC process with 40 chains and flat priori for all parameters, and run the parameter search for 1000 steps. The last 500 steps are selected for the result analysis below, which is considered to be in the equilibrium state as the chains mostly converge near 200 steps.

An advantage of using MCMC as the parameter searching tool is that it gives the extent to which a parameter can be constrained. The algorithm does not require any assumptions of the shape of the posterior PDFs, which are necessary for computing the gradient in parameter space in traditional approaches (e.g., Levenberg–Marquardt algorithm, Levenberg 1944). This is critical when the observation is barely sufficient to provide constraints as shown below, where well-constrained parameters have Gaussian-like PDFs, and poorly constrained ones have irregular, widely varying PDFs. The MCMC can also avoid convergence to local minima, which is important in this work as the relationship between the four fractal aggregate parameters are not linear, and therefore multiple local minima are expected.

3.4.4 Surface reflection

Suggested by studies of Triton's haze (Hiller et al. 1990; 1991), surface reflection may have a non-negligible contribution to the observed haze brightness. We estimate this contribution in one of the tested scenarios as shown below. In line with Hiller et al. (1990) and (1991), we use the Hapke model (Hapke 1981) to simulate Pluto's surface reflection. Due to the small optical opacity of Pluto's atmosphere, attenuation of incoming and reflected light by the atmosphere is neglected. We assume Pluto's surface scatters isotropically, so that the reflected light follows

$$I(\mu_0, \mu) = J \frac{w}{4\pi} \frac{\mu}{\mu_0 + \mu} H(\mu_0) H(\mu) \qquad (3.3)$$

where I and J are the reflected and incoming intensities, respectively; w is the single

scattering albedo of the surface; μ_0 and μ are the cosine of solar and viewing zenith angles, respectively; $H(\mu)$ is the Hapke function

$$H(\mu) = \frac{1+2\mu}{1+2\sqrt{1-w}\mu} \qquad (3.4)$$

We use the Hierarchical Equal Area isoLatitude Pixelization method (HEALPix; Górski et al. 2005) to discretize Pluto's surface for the reflection computation. This technique divides the surface into pixels with the same area and distributed uniformly on the sphere. The parameter Nside in HEALPix is set to 16, which results in a 3072-pixel map with a spatial resolution of ~4 °.

3.5 Results

Given the valuable but limited observations, degeneracy appears when the number of free parameters is too large. Therefore, we tested a number of scenarios of haze morphology in the retrieval (Table 3.2), which are compromises between fitting all the observations and limiting the degrees of freedom. The scenarios included in this work are (1) monodispersed fractal aggregates with variable dimension, (2) monodispersed fractal aggregates with variable dimension and surface reflection, (3) monodispersed fractal aggregates with variable dimension and monodispersed spheres, (4) two populations of monodispersed spheres, and

Table 3.2. Summary of scenarios tested in this work.

	Distribution	Morphology	Number of free parameters
1	Monodispersed	Aggregates with varying D_f	4
2	Monodispersed	Aggregates with varying D_f & Surface	4
3	Bimodal	Aggregates with varying D_f & Spheres (D_f=3)	6
4	Bimodal	Spheres (D_f=3) & Spheres (D_f=3)	4
5	Log-normal	Spheres (D_f=3)	3
6	Power-law	Spheres (D_f=3)	2
7	Exponential	Spheres (D_f=3)	2
8	Log-normal	Aggregates with D_f=2	3
9	Power-law	Aggregates with D_f=2	2
10	Exponential	Aggregates with D_f=2	2

(5)-(10) log-normal, power-law or exponential distribution of two-dimensional aggregates or spheres. The first three scenarios are presented in this section and the rest seven are in Appendix 3.8.3. As shown below, scenario (3), a bimodal distribution consisting of large fractal aggregates and small particles, is the only one that can fit all the observations.

3.5.1 Monodispersed fractal aggregates

Under this scenario, we assume that there is a monodispersed population of haze particles at each altitude and include four free parameters. Three of them (D_f, r_m, N_m) describe the morphology of haze particles, and the fourth one (n_a) is the aggregate local number density.

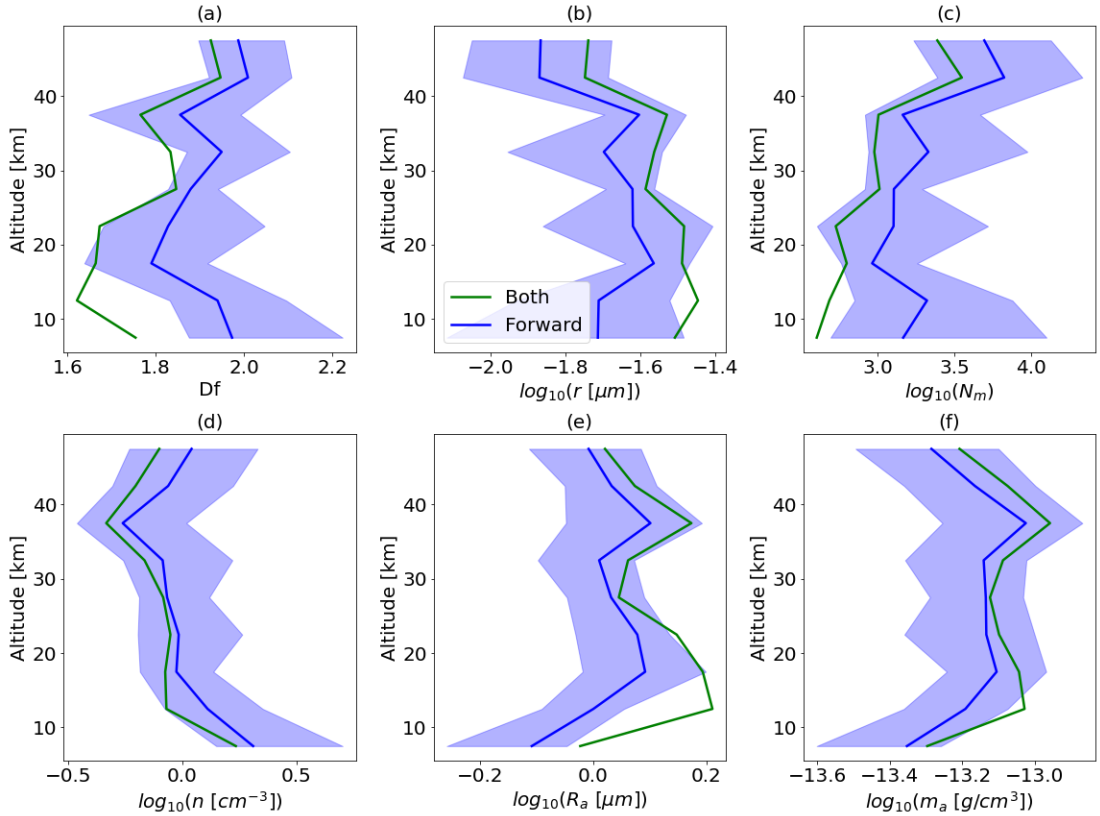

Figure 3.3. Vertical profiles of the free parameters: (a) fractal dimension, (b) monomer radius, (c) number of monomers in each aggregate, and (d) aggregate number density, and two derived quantities: (e) aggregate effective radius, and (f) total mass density assuming a material density of 1g/cm^{-3} in the retrievals under scenarios of monodispersed fractal aggregates constrained using all observations (green) and forward scattering only (blue). Results of the best-fit are denoted as solid lines, while those derived from probability density functions (PDFs) with 1-σ uncertainties are shown as shaded areas.

Vertical profiles of these four quantities that best fit the observations are shown in Figure 3.3, and the corresponding simulated observables are given in Figure 3.4. The best-fit set of parameters is that which results in the maximal posterior probability as defined in Equation (2.2). The results show that the assumption of monodispersed fractal aggregates cannot fit all of the observations: it underestimates the backward scattering (Figure 3.4a and 3.4b) due to the forward-scattering dominated phase functions of fractal aggregates. This agrees with the discrepancy suggested by Cheng et al. (2017) that one population of fractal aggregates or spheres alone cannot explain both the forward and backward scattering in the visible and the extinction in the UV.

3.5.2 Monodispersed fractal aggregates and surface reflection

As fractal aggregates tend to underestimate the backward scattering, and surface reflection usually has larger backward scattering than forward, surface reflection may fill in the gap between observed backward scattering intensity and that scattered by aggregates. Here we test the scenario where these two factors are combined. The observed intensity above Pluto's limb is assumed to consist of two parts: (1) sunlight scattered once by haze particles into New Horizon's instruments, and (2) sunlight first reflected off of Pluto's surface and then scattered once by the haze particles into New Horizon's instruments.

Given the fact that surface reflection usually contributes little when the viewing zenith angle is large, we first omit the backward scattering observables and only use the forward ones (LORRI at 148.3 °and 169.0 °, LEISA, and MVIC at 169.4 °) together with the UV extinction to constrain the haze morphology. Results of the profiles of the four free parameters (D_f, r_m, N_m, n_a) are compared to the observations in Figure 3.3 and 3.4. The simulated forward scattering of haze particles agrees with observations well. All four parameters are well-constrained across all the altitudes. The haze particles are ~1μm two-dimensional aggregates with ~20nm monomers at most of the altitudes (Figure 3.3b and 3.3e). Phase functions of these haze particles are forward scattering dominated, which explains the intense forward scattering observed. On the other hand, backward scattering by these particles is

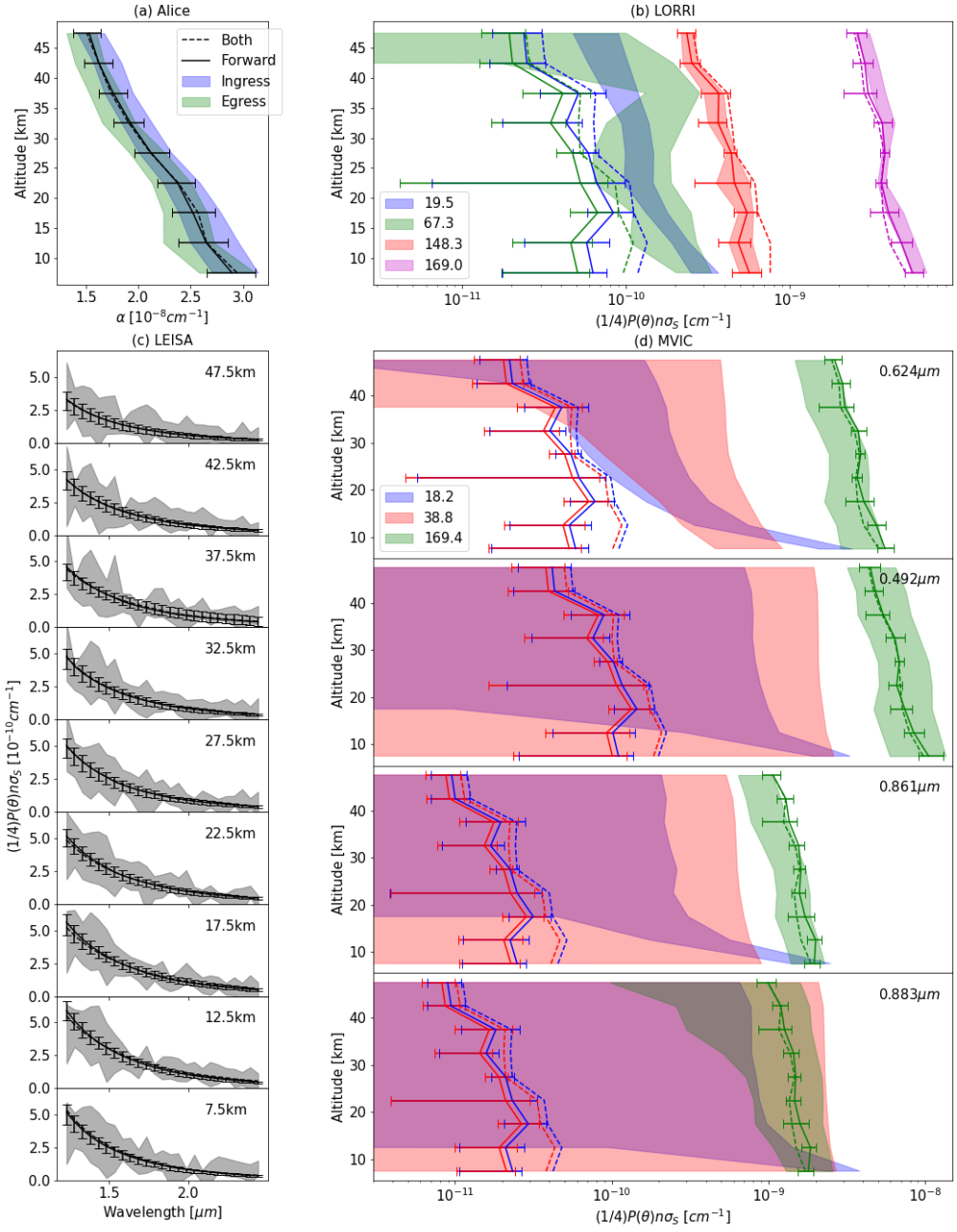

Figure 3.4. Same as Figure 3.1, but the model results are under scenarios of monodispersed fractal aggregates. The best-fit of the results constrained using all observations are denoted as dashed lines, and the results with 1-σ uncertainties constrained using forward scattering only are shown as solid lines with error bars.

underestimated. The local scattering intensity of haze particles is less than the observed value by a factor of two (Figure 3.4b), which results in a LOS I/F difference of ~$5*10^{-3}$ in the lower 50km. Therefore, we test the inclusion of surface reflection to try to fill this gap.

The secondary scattering of light reflected off of Pluto's surface by haze particles is computed using our retrieved haze morphology. We estimate an upper limit by computing the I/F contribution from this source when the incident and emission vectors are in the same specular plane. For each discretized point along the LOS, the secondary scattering is computed by summing up the reflected light from the pixelated Pluto surface multiplied by the haze scattering function with the corresponding geometry; integrating the secondary scattering along the LOS then provides the observables. Results for various surface single scattering albedos are shown in Figure 3.5, suggesting that even with the highest albedo (w=1), the maximum value of the LOS-integrated secondary scattering, which is around 90 °

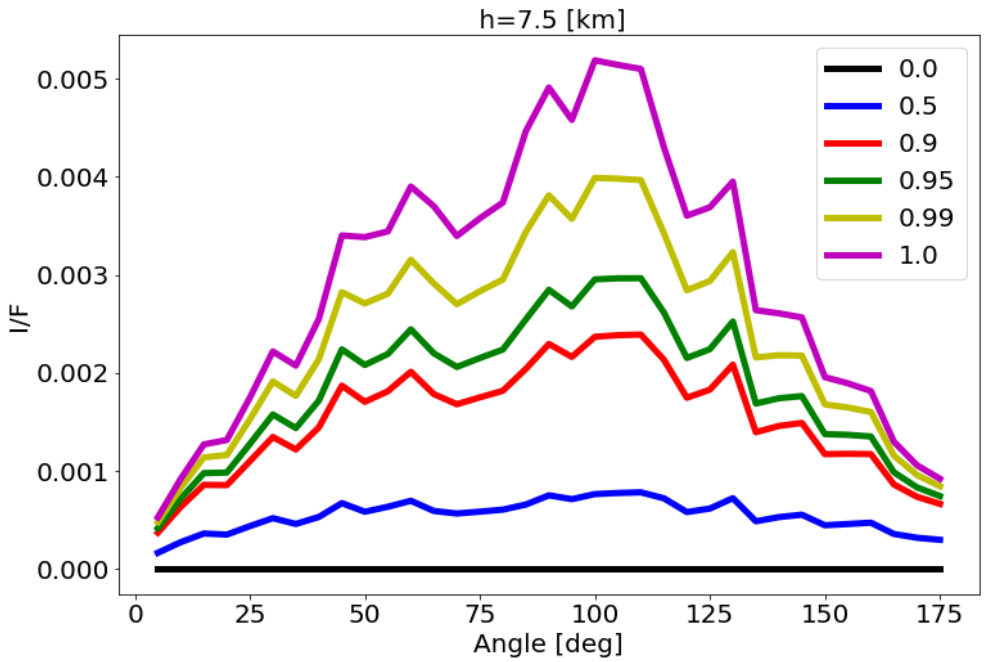

Figure 3.5. I/F of secondary scattering by Pluto's surface and then by haze particles integrated along line-of-sight (LOS) at an altitude of 7.5km under assumptions of surface material single scattering albedos of 0.0 (black), 0.5 (blue), 0.9 (red), 0.95 (green), 0.99 (yellow), and 1.0 (magenta).

phase angle, is smaller than half of the gap between the primary scattering and observed I/F. Moreover, the maxima at the observed phase angles of LORRI (19.5 °, 67.3 °, 148.3 °, and 169.0 °) are around or less than $2.5*10^{-3}$, so the upper limit of the secondary scattering is about one order of magnitude lower than the required backward scattering.

3.5.3 Bimodal distribution

As monodispersed fractal aggregates and surface reflection cannot reproduce the observed backward scattering, we consider a combination of a bimodal distribution of two haze particle populations with different sizes. Besides the aggregates as described in scenario (1), we include another population in the retrieval. The second population is assumed to be spheres, but the fractal dimension does not matter as will be discussed in Section 3.6. The

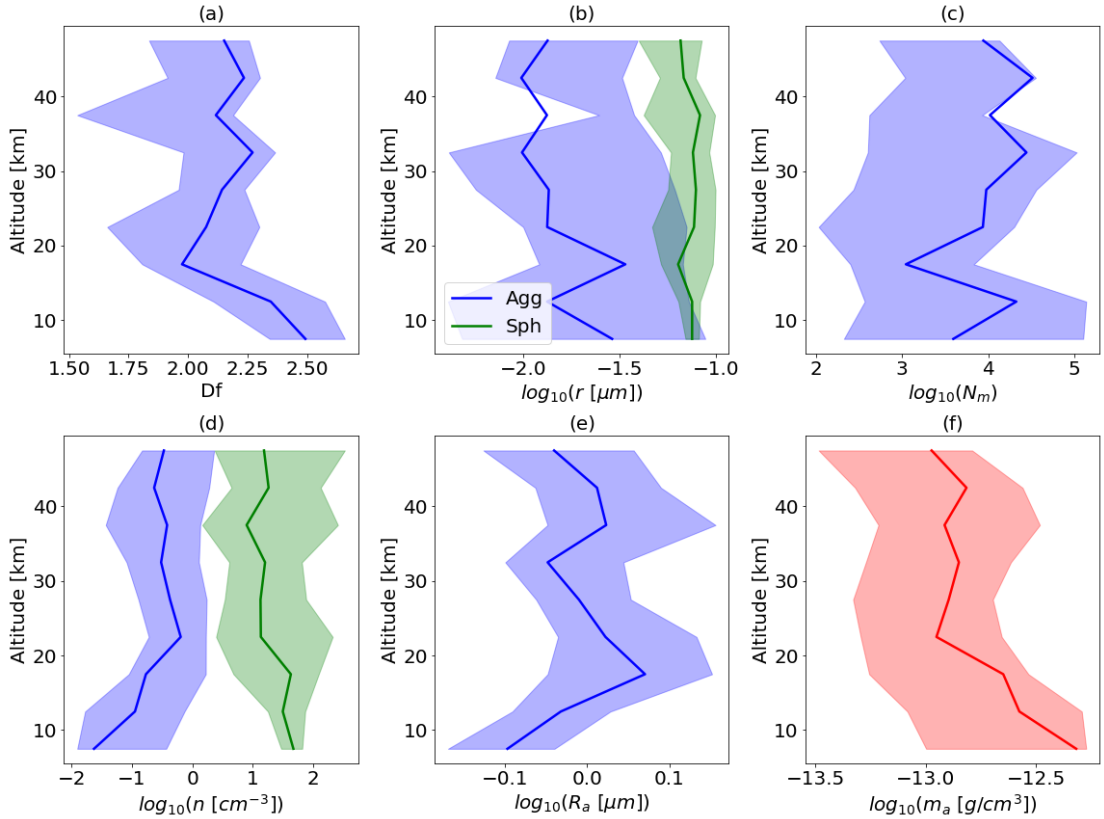

Figure 3.6. Same as Figure 3.3, but for the scenario of a bimodal distribution of large fractal aggregates and small spheres. Parameters of fractal aggregates are denoted as blue, while those of spheres are shown as green. The total mass density in (f) consists of both aggregates and spheres.

spheres are parameterized with two variables, radius (r_s) and number density (n_s). Therefore, six free parameters in total are considered in the retrieval. After searching for the parameters using MCMC, vertical profiles of these six parameters are shown in Figure 3.6, and their PDFs at one of the altitudes are given in Figure 3.7. Similar to those in scenario (2), haze particles are ~1μm two-dimensional aggregates with ~20nm monomers across the altitude region, while the spheres have radii ~80nm. These two radii in the bimodal distribution are similar to those simulated in Titan's atmosphere, which agrees well with observations (Lavvas et al. 2010). Comparison between the modeled observables and

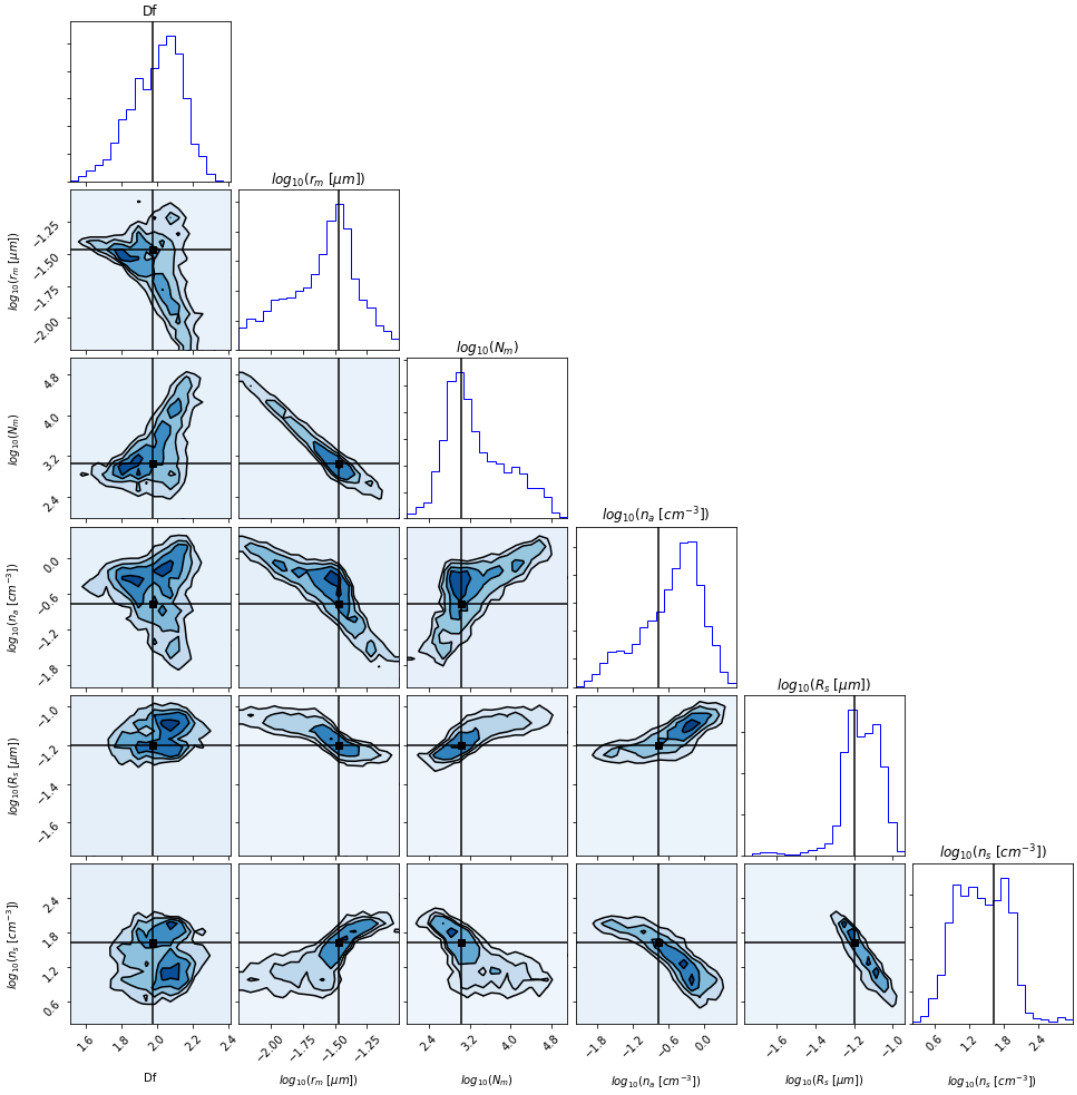

Figure 3.7. Probability density functions (PDFs) of the six free parameters and their joint distribution of the retrieval at 22.5km under the scenario of the bimodal distribution of aggregates and spheres. The values of the best-fit are denoted as solid black lines.

observations (Figure 3.1) indicates that under this scenario both the large forward and backward scattering and the UV extinction could be explained. The two types of haze particles have comparable UV extinctions with each dominating one of the forward and backward scattering at visible wavelengths (Figure 3.8).

3.6 Discussion

We investigate a number of scenarios in the retrieval, and conclude that the observed large forward and backward scattering intensities at visible wavelengths and UV extinction likely stem from a bimodal distribution of haze particles. The forward scattering observations of LEISA constrain the large population's effective radius to ~1μm, while small particles (~80nm) dominate the backward scattering and contribute to approximately half of the UV

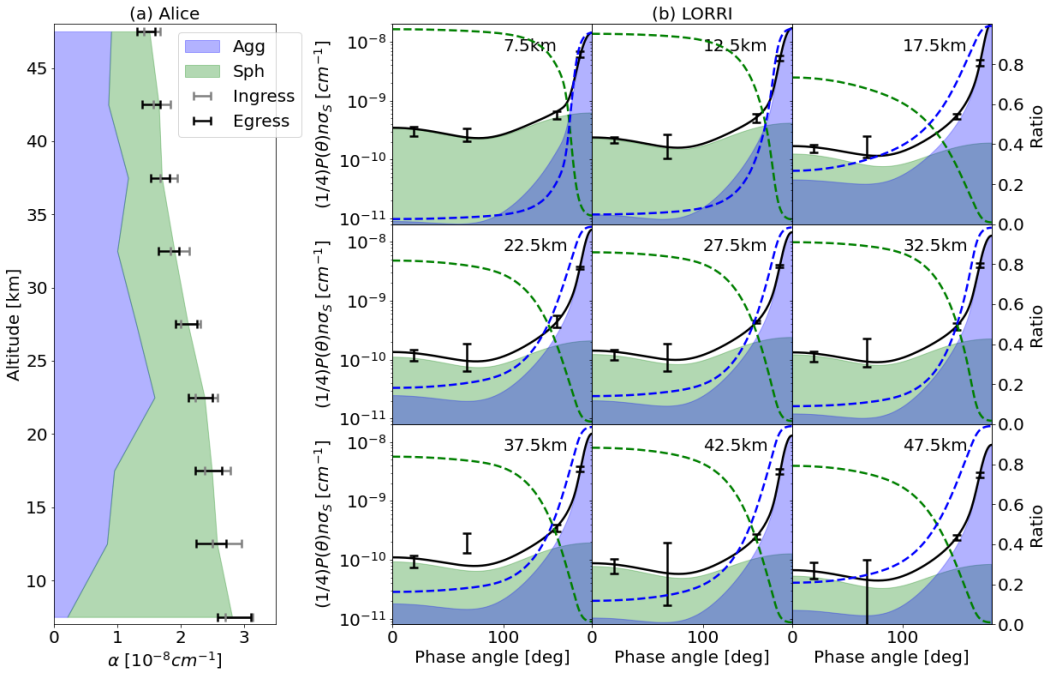

Figure 3.8. Contributions of fractal aggregates (blue) and spheres (green) to (a) UV extinction and (b) scattering in visible. (a) Observations obtained by Alice spectrograph during solar occultation are denoted during ingress and egress are denoted as grey and black error bars. (b) Observations obtained by LORRI are denoted as black error bars. The solid black line is the total contribution of aggregates and spheres, which is also the sum of the shaded areas. The colored dashed lines represent the ratio of contribution from each component.

extinction (Figure 3.8a). Medium size particles are ruled out due to the fact that they do not possess strong forward and backward scattering with the same UV extinction. Therefore, power-law or exponential distributions cannot satisfy the observations. Our results break the degeneracy between bimodal and power-law distributions of haze particles suggested by Kutsop et al. (in review), who used only MVIC observations.

A bimodal distribution of aerosols has also been found and simulated in Titan's atmosphere using observations from Cassini/Huygens (Tomasko et al. 2008, Lavvas et al. 2010). Lavvas et al. (2010) successfully modeled it by including a high altitude monomer production region (~1000km above Titan's surface) and a transition of aggregate dimension from spheres to two-dimensional aggregates near 520km. Due to different collision cross-sections and transport velocities of these two types of particles, the bimodal distribution forms near the transition region. The ultimate sizes of these two populations in Titan's atmosphere (3.4μm and 84nm) are similar to those found in this work. Therefore, there may be a dimension transition in Pluto's atmosphere as well. Another possible mechanism is condensation/sticking of gaseous species. Due to the low temperature, C_2 species may condense/stick onto haze particles through gas/solid phase change (Wong et al. 2017) or gas-haze incorporation (Luspay-Kuti et al. 2017). Proof or negation of these mechanisms requires further modeling investigations and is beyond the scope of this work.

The fractal dimension of the small-particle population is assumed to be three in our retrieval for the purpose of simplicity and for reducing the number of free parameters. However, it cannot be constrained with currently available observations. The difference in scattering between a three-dimensional sphere and a two-dimensional aggregate with the same effective radius of 80nm is less than a factor of three (Figure 3.9), which is less than the observational uncertainty. Tighter constraints on the dimension of the small-particle population require scattering phase function observations at much shorter wavelengths that are comparable to the particle radius.

The vertical profiles of the retrieved parameters show a possible change in the dimension of the aggregates below 15km (Figure 3.6a), while the agreement with the observations

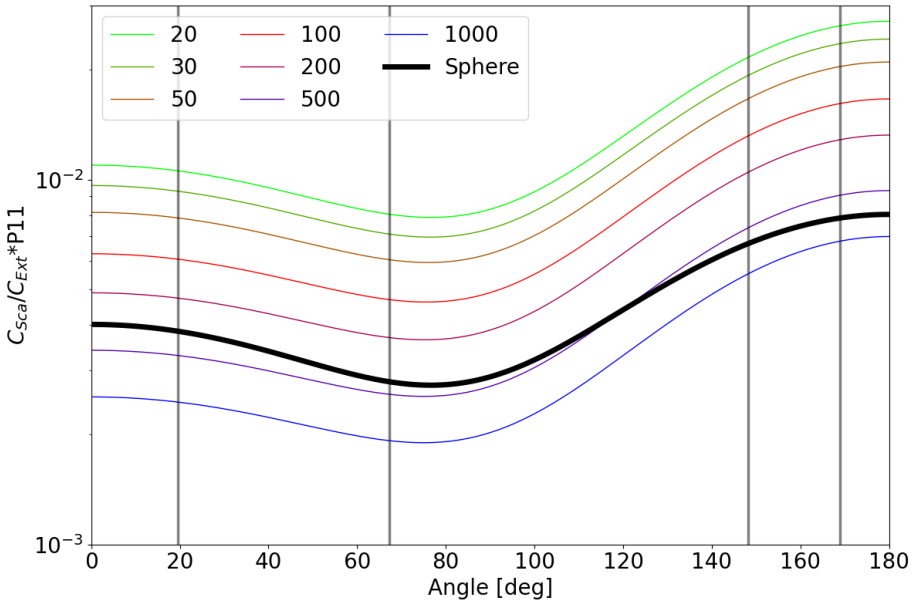

Figure 3.9. Local scattering intensity with unit UV extinction of 80nm spheres (black solid line) and two-dimensional aggregates with the same effective radius but different number of monomers (colored solid lines). The four phase angles 19.5 °, 67.3 °, 148.3 °, and 169.0 °of LORRI observations are shown as vertical solid grey lines.

simultaneously deteriorates (Figure 3.1b and 3.1d) as the observed backward scattering I/F increases rapidly towards the surface. This may come from some boundary effects, e.g. aerosols lifted by wind or some rapid change of atmospheric properties near the surface. In either case, more degrees of freedom need to be considered at these near-surface regions but the current limited observations may not be able to provide sufficient constraints.

While the contribution of Pluto's surface reflection is negligible on the observed I/F above Pluto's limb, the retrieved haze optical properties can help improve the constraints on Pluto's surface albedo (e.g., Buratti et al. 2017). Radiative transfer models (e.g., Spurr & Natraj 2011) can also be deployed into this context. Constraints on the composition of Pluto's surface may be considerably improved if the influence from the aerosols is removed.

3.7 Summary

We find that Pluto's haze particles follow a bimodal distribution in the lower 50km through a joint retrieval of available observations obtained by multiple instruments onboard the New

Horizons spacecraft. The bimodal distribution consists of a population of ~1μm two-dimensional aggregates with ~20nm monomers and a population of ~80nm particles. The fractal dimension of the small-particle population cannot be constrained using current observations. This distribution may come from a transition of haze particle fractal dimension in Pluto's atmosphere. Surface reflection provides negligible contributions to haze brightness above the limb, but other surface-related mechanisms may affect the haze brightness in the lower 15km of the atmosphere.

3.8 Appendix

3.8.1 Abel transform of noisy data

In remote sensing of planetary atmospheres, observables are usually line-of-sight (LOS) integrated quantities, necessitating it to convert these observables to local quantities for analysis of physical or chemical properties of the atmosphere. In the solar occultation by Alice, the observable is LOS optical depth.

$$\tau_{LOS}(r) = \int_{-\infty}^{+\infty} n(s)\sigma_{ext}(s)ds \qquad (3.S1)$$

where τ_{LOS} is the LOS optical depth; r is the distance of Pluto's center to LOS; n is the local number density; σ_{ext} is the extinction cross section; s is the path along LOS. In the limb scattering geometry of the other three instruments, the observed I/F is an integration of local scattering intensity.

$$I/F\ (r) = \int_{-\infty}^{+\infty} \frac{1}{4} P(cos\theta) n(s)\sigma_{sca}(s)\ ds \qquad (3.S2)$$

where σ_{sca} is the scattering cross section; θ is the scattering angle; P is the phase function, which is normalized to 4π. Given the same form of Equation (3.S1) and (3.S2), they can be unified as follows:

$$N(r) = \int_{-\infty}^{+\infty} D(s)\, ds \qquad (3.S3)$$

where N is the observable; D is the corresponding local quantity. Assuming a spherical asymmetry, D is a function of r, and the equation becomes an Abel integral.

$$N(r) = 2\int_{r}^{+\infty} D(r')\frac{r'}{\sqrt{r'^2 - r^2}}\, dr' \qquad (3.S4)$$

where r' is the distance of Pluto's center to the point of each D in the integral. This equation shows that N only depends on D at altitudes higher than the impact parameter. Abel inverse transform suggests an exact solution (Roble and Hays, 1972).

$$D(r') = -\frac{1}{\pi}\int_{r'}^{+\infty} \frac{dN(r)/dr}{\sqrt{r^2 - r'^2}}\, dr \qquad (3.S5)$$

However, the exact solution is not a good option in the inversion of observations with noisy data, as the derivative of N in the integral is sensitive to noise. Therefore, instead of using the exact solution, we rewrite Equation (3.S4) with discrete altitude bins and solve the problem through linear regression.

$$N_i = 2\sum_{j=1}^{n} D_j \int_{r_{1j}}^{r_{2j}} \frac{r'}{\sqrt{r'^2 - r_i^2}}\, dr' \qquad (3.S6)$$

where i and j are the indices of altitude bins; n is the total number of bins; r_{1j} and r_{2j} are the lower and upper boundary of the j-th altitude bin, and $r_{1(j+1)} = r_{2j}$; N_i and D_j are the corresponding LOS and local quantities at the i-th and the j-th altitude bins, respectively. In this case, all the integrals form a geometry matrix **A**, which consists of element A_{ij} for each pair of (i, j).

$$A_{ij} = \int_{r_{1j}}^{r_{2j}} \frac{r'}{\sqrt{r'^2 - r_i^2}}\, dr' \qquad (3.S7)$$

A is an upper triangular matrix as $A_{ij}=0$ when i>j. Including the observation noise, there is a linear relationship between vectors of N and D.

$$\boldsymbol{N} = \boldsymbol{A} \cdot \boldsymbol{D} + \boldsymbol{\varepsilon} \qquad (3.S8)$$

where **N** and **D** are column vectors which have N_i and D_j as the elements; **ε** is the vector with observation noise for each corresponding N_i. Then the linear problem can be solved as follows:

$$\boldsymbol{K} = (\boldsymbol{A}^T \boldsymbol{C_N} \boldsymbol{A})^{-1} \cdot (\boldsymbol{A}^T \boldsymbol{C_N^{-1}}) \qquad (3.S9)$$

$$\boldsymbol{D} = \boldsymbol{K} \cdot \boldsymbol{N} \qquad (3.S10)$$

$$\boldsymbol{C_D} = \boldsymbol{K} \cdot \boldsymbol{C_N} \cdot \boldsymbol{K}^T \qquad (3.S11)$$

where $\mathbf{C_D}$ is the covariance matrix of **D**; $\mathbf{C_N}$ is the covariance matrix of **N**. Square root of the diagonal elements in $\mathbf{C_D}$ serve as the uncertainties of local densities. Inversion through linear regression does not contain derivatives, so it is more robust against observation noise. Further regularization, which is not included in this work as the result is already acceptable, can also be added to decrease the influence of noise (Quémerais et al. 2006).

The transform requires the upper boundary of the altitude bins to be sufficiently high so that the truncation from infinity to r_{2n} in Equation (3.S6) can be neglected. However, some of the observations have limited altitude range due to the instrument field of view. The I/F profile obtained by LORRI at phase angle of 67.3 °is limited to the lower 50km, which is the smallest altitude range among all observations, and therefore constrains the altitude range that we can conduct the analysis. Given the fact that Pluto's haze extends up to ~200km above the surface with a densityscale height of ~50km (Gladstone et al. 2016), haze above the available observation range is not negligible. Due to the small radius of Pluto (1190km, Hinson et al. 2017), the change of scale height with gravity needs to be considered in the extrapolation. We assume an exponential decay in geopotential of the local density, in line with that in Young et al. (2018). The first order approximation is as follows (Young 2009):

$$H = H_0 \frac{r^2}{r_0^2} \qquad (3.S12)$$

$$D = D_0 e^{-\frac{r_0}{H_0}\left(1-\frac{r_0}{r}\right)} \qquad (3.S13)$$

$$N = N_0 e^{-\frac{r_0}{H_0}\left(1-\frac{r_0}{r}\right)} \left(\frac{r}{r_0}\right)^{\frac{3}{2}} \frac{1+\frac{9}{8}\frac{H}{r}}{1+\frac{9}{8}\frac{H_0}{r_0}} \qquad (3.S14)$$

where H is the scale height; the subscript 0 indicates the surface value of corresponding variables. Through this approach, we extrapolate N to 2000km above Pluto's surface by fitting Equation (3.S14) to observables. The ranges selected for the fitting are between 25km to the highest valid altitude, 25–75km and 15–50km for LORRI, LEISA and MVIC observations.

The extrapolation results in two parts for the altitude bins, which requires Equation (3.S8) to be rewritten.

$$\begin{bmatrix} \boldsymbol{N}_1 \\ \boldsymbol{N}_2 \end{bmatrix} = \begin{bmatrix} \boldsymbol{A}_{11} & \boldsymbol{A}_{12} \\ 0 & \boldsymbol{A}_{22} \end{bmatrix} \cdot \begin{bmatrix} \boldsymbol{D}_1 \\ \boldsymbol{D}_2 \end{bmatrix} + \begin{bmatrix} \boldsymbol{\varepsilon} \\ 0 \end{bmatrix} \qquad (3.S15)$$

where subscript 1 denotes the corresponding variables at altitudes with valid observations, while subscript 2 denotes those of extrapolation. We assume the extrapolation to be precise and neglect its uncertainty. Combining Equation (3.S15) with (3.S9-3.S11), we can finally derive the required local density and its uncertainty.

$$\boldsymbol{D}_2 = (\boldsymbol{A}_{22}^T \boldsymbol{A}_{22})^{-1} \cdot (\boldsymbol{A}_{22}^T \boldsymbol{N}_2) \qquad (3.S16)$$

$$\boldsymbol{K}_1 = \left(\boldsymbol{A}_{11}^T \boldsymbol{C}_{\boldsymbol{N}_1} \boldsymbol{A}_{11}\right)^{-1} \cdot \left(\boldsymbol{A}_{11}^T \boldsymbol{C}_{\boldsymbol{N}_1}^{-1}\right) \qquad (3.S17)$$

$$\boldsymbol{D}_1 = \boldsymbol{K}_1 \cdot (\boldsymbol{N}_1 - \boldsymbol{A}_{12} \boldsymbol{D}_2) \qquad (3.S18)$$

$$\boldsymbol{C}_{\boldsymbol{D}_1} = \boldsymbol{K}_1 \cdot \boldsymbol{C}_{\boldsymbol{N}_1} \cdot \boldsymbol{K}_1^T \qquad (3.S19)$$

where $\mathbf{C_{D1}}$ and $\mathbf{C_{N1}}$ are the covariance matrices of $\mathbf{D}_1$ and $\mathbf{N}_1$, respectively. $\mathbf{D}_1$ corresponds to the UV extinction coefficient, $n\sigma_{ext}$, or local scattering intensity, $\frac{1}{4}Pn\sigma_{sca}$ in the integral of Equation (3.S1) and (3.S2), respectively.

3.8.2 MVIC data processing

Table 3.S1. List of MVIC observations

MET	Time	Phase Angle (degree)	Distance (10^3km)	Resolution (km)
0299162512	2015-07-14 T06:50:12	18.2	246.1	4.87
0299178092	2015-07-14 T11:10:52	38.8	33.1	0.67
0299193157	2015-07-14 T15:20:58	169.4	175.3	3.47

Details of the MVIC instrument design and operation are given in Reuter et al. (2008). Science operation of MVIC is under TDI mode with two panchromatic and four color arrays. We select the observations obtained by color arrays as they contain Pluto haze's spectral characteristics. Level 2 data is used for analysis in this work, whose identifier on NASA PDS is *New Horizons MVIC Pluto Encounter Calibrated Data v3.0* (Stern, 2018). The data contains bias-subtracted, flattened images, but it does not include corrections for scattered light, cosmic rays, geometric and motion distortion. We follow the *New Horizons SOC to Instrument Pipeline ICD* (NASA PDS) to convert the calibrated data number (DN) to I/F, then use the SPICE system (Acton 1996) to compute the geometry with navigation data (Figure 3.S1). With the geometry information, we determine the resolution and mean phase angle of each image. Three observations with resolution better than 5km/pixel are selected (Table 3.S1). As the geometry inferred by the navigation data does not perfectly locate Pluto in each image (Figure 3.S1), we conduct a further correction as the haze altitude needs to be accurate. A one pixel offset can result in an altitude difference as large as 5km, which is the bin interval selected in this work. To correct the Pluto location in each image, we select the points with the largest brightness gradient around the sun-lit edge (Figure 3.S1), then fit a circle to these points, which serve as the Pluto's edge. Other geometries are offset according

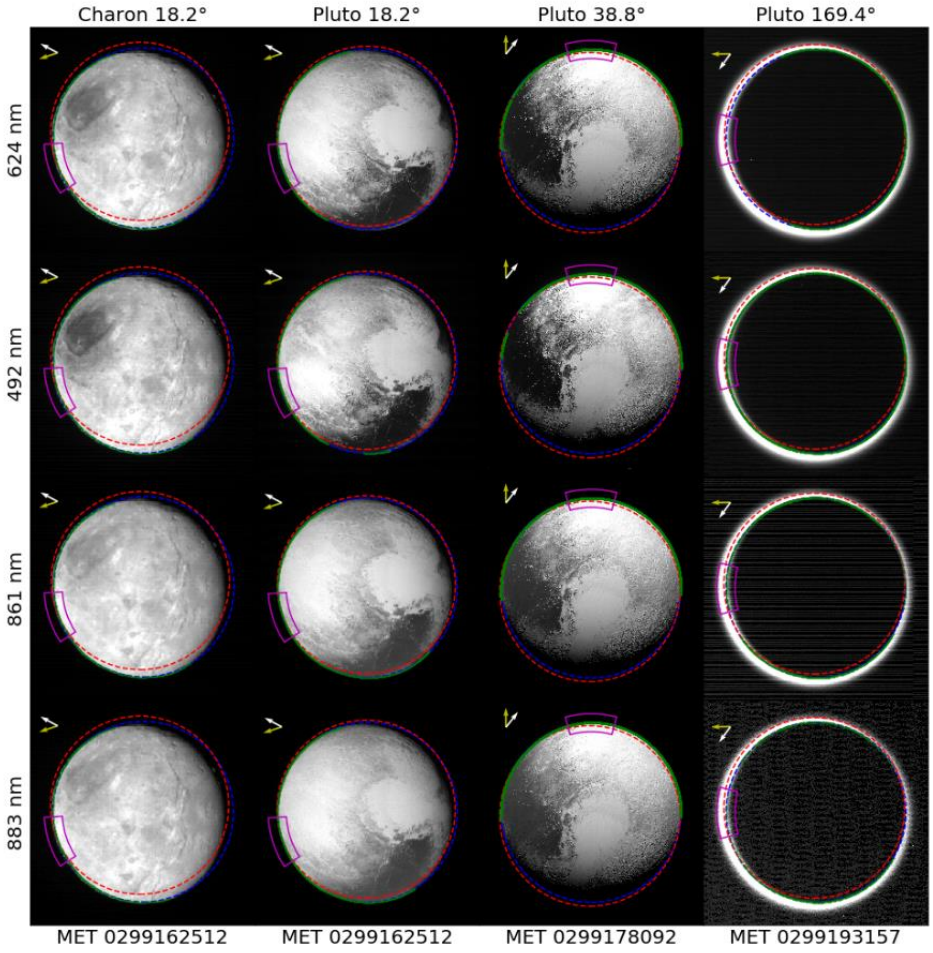

Figure 3.S1. MVIC images of Charon and Pluto at mission elapsed time (MET) of 0299162512, 0299178092, and 0299193157. Images at different channels are presented at each row, and the columns are for target and phase angles. Red and blue dashed circles denote the prediction of target locations and those derived by fitting the edge of targets using the green dots, respectively. The magenta curved boxes show the regions which are used for data analysis and presented in Figure S2. The yellow and which arrows denote the direction of the Sun and Pluto's north.

to the correction. Pluto is not a perfect sphere in the images, which may be due to the shape itself or the fast motion of the spacecraft at small distances. We select regions with the sharpest edge, as haze altitudes in these regions are computed with the best accuracy (Figure 3.S1).

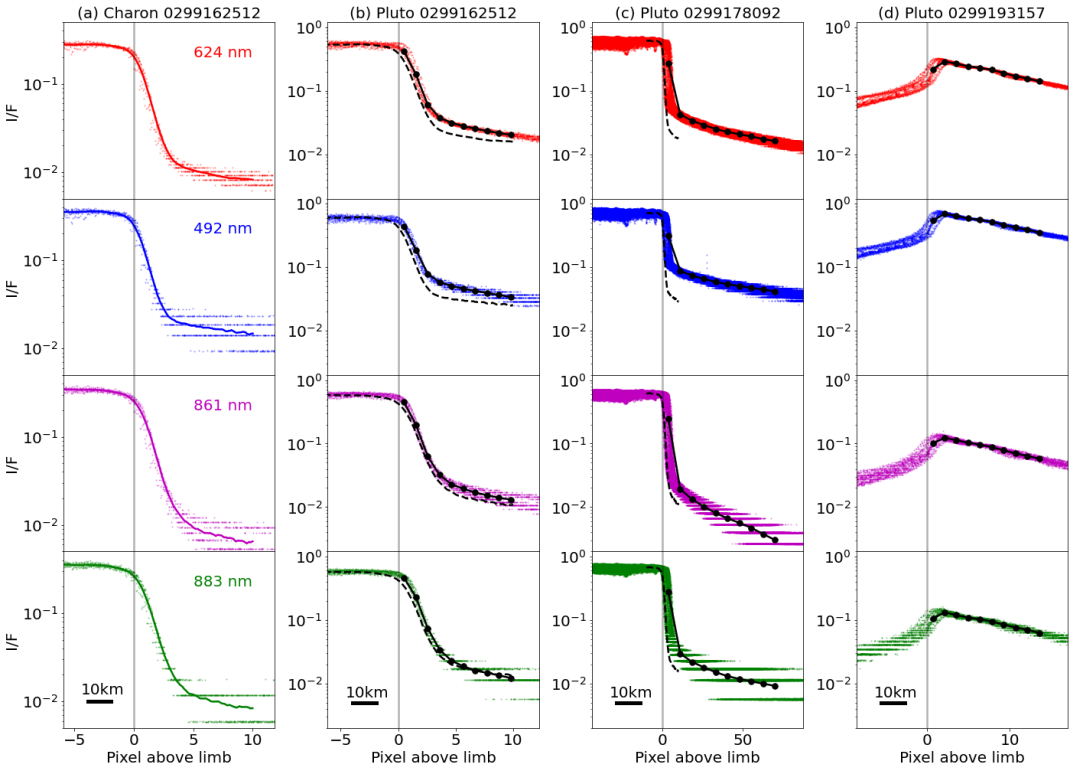

Figure 3.S2. I/F profiles as a function of pixel distance above the limn. Sequence of the panels are the same as Figure 3.S1. Observed values are denoted as colored dots, and the vertical solid grey lines show the location of target limbs. Moving average of observations of Charon are shown as colored solid lines in (a), and used for stray light correction shown as black dashed lines in (b) and (c). Averaged observations of Pluto are denoted as black dots and solid lines in (b)-(d). A black bar showing the scale of 10km is at the lower left of each column.

Stray light is not negligible in the images, similar to that in LORRI observations (Cheng et al. 2017). It is especially important at backscattering configuration when the disk brightness is much higher than the haze. As the stray light is an instrument effect, it is a function of pixel distance above the limb of celestial bodies, which can also be seen around the airless Charon (Figure 3.S2). I/F profiles of Pluto's haze in LORRI observation are corrected by subtracting the normalized stray light above Charon's limb (Cheng et al. 2017). Here, we include the correction the same way. We compute the moving averages of I/F profiles as a function of pixel distance from Charon's limb, then subtract them from Pluto's. As Charon appears in only one of the observations (MET 0299162512), and the instrument effect should stay the same during the flyby, we assume the stay light influence is the same among images obtained

by the same color array at different time, and apply the correction individually for each array (Figure 3.S2). The I/F profiles after this correction are used in the haze property analysis.

3.8.3 Test of scenarios

Besides the three scenarios presented in section 3.5, the rest seven are tested here to fit observations: (4) bimodal distribution of spheres (haze particles with D_f=3), (5) log-normal distribution of spheres, (6) power-law distribution of spheres, (7) exponential distribution of spheres, (8) log-normal distribution of two-dimensional aggregates (haze particles with D_f=2), (9) power-law distribution of two-dimensional aggregates, and (10) exponential distribution of two-dimensional aggregates. A comparison of the goodness among the scenarios, including those mentioned in section 3.5, is given in Figure 3.S3. We quantify the goodness using the maximal probability that can be reached under each scenario as defined in Equation (3.2). All of the seven scenarios show the goodness far worse than the bimodal distribution of large aggregates and small particles.

Scenario (4) of a bimodal distribution of spheres contains four free parameters. They are the sphere radii and the two corresponding number densities. Scenario (5) contains three free parameters, which are the two parameters defining the log-normal distribution and the total number density. We include 30 size bins of spheres, covering a radius range from 1.3nm to 1.0μm. Each bin is assumed to have particles that contains twice of the monomers of the previous, so the ratio between the radii between two consecutive bins is $\sqrt[3]{2}$. Therefore, the number density of particles in each bin is defined as follows.

$$n_i = n_0\big(CDF(r_{2i}) - CDF(r_{1i})\big) \qquad (3.S20)$$

where i is the bin index; r, r_{1i}, r_{2i} are the radii at the bin center, lower and upper boundary of the i-th bin, respectively; n_i is the number density of particles in the i-th bin; n_0 is the total number density. CDF is the cumulative density function, which is defined as below for the log-normal distribution.

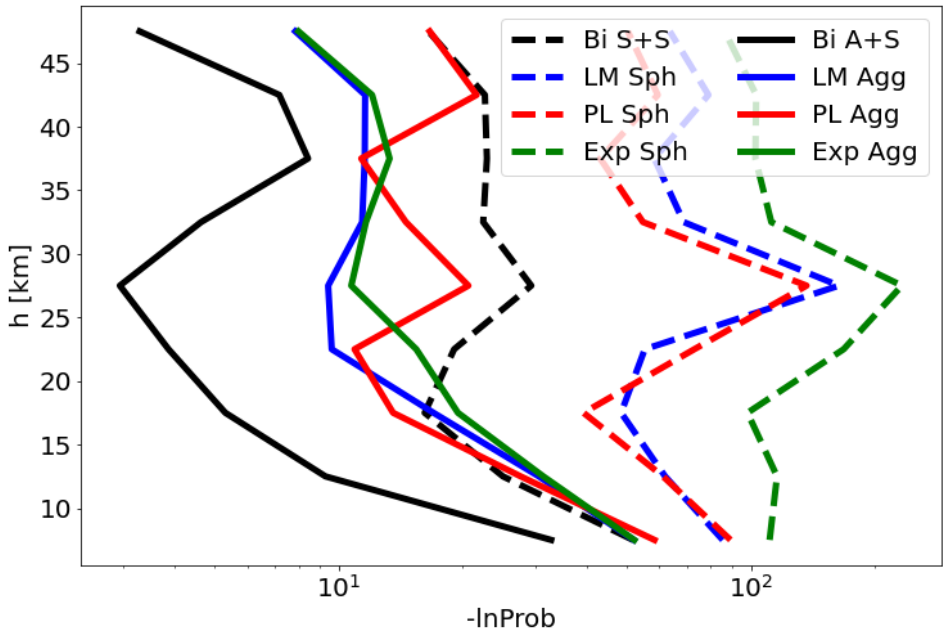

Figure 3.S3. Comparison of the goodness of scenarios, quantified by the negative of the logarithm of the posterior probability: bimodal distribution of spheres (black dashed line), log-normal distribution of spheres (blue dashed line), power-law distribution of spheres (red dashed line), exponential distribution of spheres (green dashed line), bimodal distribution of aggregates and spheres (black solid line), log-normal distribution of aggregates (blue solid line), power-law distribution of aggregates (red solid line), and exponential distribution of aggregates (green solid line).

$$CDF_{LN}(r) = \frac{1}{2} + \frac{1}{2}\, erf\left(\frac{ln\, ln\,(r)\, - \mu}{\sqrt{2}\sigma}\right) \qquad (3.S21)$$

where r is the particle radius; μ and σ are the mean and standard deviation of the logarithm of radius, respectively; erf is the error function. Under this scenario, n_0, μ and σ are the three free parameters. Scenario (6) is the same as scenario (5) except for the CDF, which is defined as below.

$$CDF_{PL}(r) = 1 - r^{1-p} \qquad (3.S22)$$

where p is the power describing how fast the number density decreases with size. As the CDF is defined using one parameter, this scenario has two free parameters (n_0 and p).

Similarly, scenario (7) has the CDF defined as below.

$$CDF_{Exp}(r) = 1 - e^{-\alpha r} \qquad (3.S23)$$

where α is the exponent describing the decrease of number density with size. This scenario also has two free parameters (n_0 and α).

Scenarios (8)-(10) are the same with scenarios (5)-(7), respectively, except for the particle bins. A monomer size (r_m) is assumed and fixed during each retrieval, and we tested a group of retrievals with monomer size from 1nm to 0.1μm. Monomer number ratio is still 2 between consecutive bins. As the fractal dimension is fixed to be 2, the effective radius ratio is $\sqrt{2}$. The smallest size bin contains 2 monomers, and the largest contains 2^{30}, which result in an effective radius ~$3*10^5$ times of that of the monomer. The optimal monomer sizes to reach the maximal probabilities are 20, 20, and 30nm under the scenarios (8)-(10), respectively. The comparison of the corresponding goodness is shown in Figure 3.S3.

Chapter 4

Earth as an Exoplanet: A Two-dimensional Alien Map

Siteng Fan[1], Cheng Li[1], Jia-Zheng Li[1], Stuart Bartlett[1],

Jonathan H. Jiang[2], Vijay Natraj[2], David Crisp[2], Yuk L. Yung[1,2]

[1]California Institute of Technology, Pasadena, CA 91125

[2]Jet Propulsion Laboratory, California Institute of Technology, Pasadena, CA 91109

Fan, et al. (2019). "Earth as an Exoplanet: A Two-dimensional Alien Map". In: *Astrophysical Journal Letters* 882, L1. doi: 10.3847/2041-8213/ab3a49.

4.1 Abstract

Resolving spatially-varying exoplanet features from single-point light curves is essential for determining whether Earth-like worlds harbor geological features and/or climate systems that influence habitability. To evaluate the feasibility and requirements of this spatial feature resolving problem, we present an analysis of multi-wavelength single-point light curves of Earth, where it plays the role of a proxy exoplanet. Here, ~10,000 DSCOVR/EPIC frames collected over a two-year period were integrated over the Earth's disk to yield a spectrally-dependent point source and analyzed using singular value decomposition. We found that, between the two dominant principal components (PCs), the second PC contains surface-related features of the planet, while the first PC mainly includes cloud information. We present the first two-dimensional (2D) surface map of Earth reconstructed from light curve observations without any assumptions of its spectral properties. This study serves as a baseline for reconstructing the surface features of Earth-like exoplanets in the future.

4.2 Introduction

Since the first exoplanet was detected (Campbell et al. 1988), approximately four thousand more have been confirmed (NASA Exoplanet Archive, exoplanetarchive.ipac.caltech.edu). Among these exoplanets, a number of them have similar properties to Earth, and may be habitable, e.g. TRAPPIST-1e (Gillon et al. 2017). However, existing measurements are still not adequate to determine whether these planets can support life. A geological and climate system that supports all three phases of water is critical to life on Earth's surface. The presence of atmospheric water vapor, clouds, and surface oceans could therefore serve as biosignatures that can be observed from a distance, and are also among the indicators for habitability. Identifying surface features and clouds on exoplanets is thus essential in this context.

Earth is the only known planet that harbors life. Remote sensing observations of Earth can therefore serve as proxies for a habitable exoplanet, as seen from the perspective of hypothetical distant observers. A number of such studies have been performed since the first

analysis of snapshots of Earth obtained by the Galileo spacecraft (Sagan et al. 1993, Geissler et al. 1995). Two observations of Earth's light curve, each spanning one day, obtained by the Deep Impact spacecraft, were used to identify changes of surface features (land/ocean) and clouds (Cowan et al. 2009, 2011; Cowan & Strait 2013). Using principal component analysis (PCA), time series of disk-integrated spectra of Earth were decomposed into two dominant "eigencolors", which contained 98% of the light curve variance. Land/ocean changes and cloud patterns were recently extracted from Earth's light curves using two years of observations of the Earth's bright side from the Deep Space Climate Observatory (DSCOVR, Jiang et al. 2018). Besides the Earth, other solar system planets (e.g. Jupiter, Ge et al. 2019) have also been treated as proxy exoplanets, with their light curves analyzed to provide baselines for exoplanet studies.

Despite interference from clouds, two-dimensional (2D) surface maps of exoplanet surfaces can be constructed using time-resolved spectra together with orbital and viewing geometry information, which in principle can be derived from light curves and other observables. Spatial maps of hot Jupiter atmospheres have been constructed (e.g. Knutson et al. 2007, Louden & Wheatley 2015). However, the detection and mapping of potentially habitable Earth-like exoplanets remains a challenge, especially when surface features and clouds contribute strongly to light curves. Numerical models that simulate light curve observations of the Earth using known spectra of different surface types have been used to test methods for retrieving 2D maps (e.g. Fujii & Kawahara 2012, Cowan & Fujii 2017, Farr et al. 2018). Using two-day single-point light curves from NASA's EPOXI mission, Cowan et al. (2009) presented the first retrieved longitudinal surface map of Earth's surface. However, 2D surface maps have not yet been derived from actual single-point light curve observations, due to their low temporal or spatial resolutions.

In this paper, we reanalyze the two-year DSCOVR/EPIC observations presented by Jiang et al. (2018) to study the Earth as a proxy exoplanet. We integrate over the disk of the Earth to reduce each image to a single point source, in order to simulate the light curve of a distant exoplanet. We report the first 2D surface map of this proxy cloudy exoplanet reconstructed

from its single-point light curves, without making any assumptions about its spectral features.

4.3 Observations and labels

Disk-integrated light curves analyzed in this work are derived from Earth's images obtained by DSCOVR's Earth Polychromatic Imaging Camera (EPIC; www.nesdis.noaa.gov/DSCOVR/spacecraft.html) during 2016 and 2017. The DSCOVR spacecraft is positioned at the first Sun-Earth Lagrangian point (L1), viewing the sunlit face of Earth from a distance of about 10^6 km. From this vantage point, DSCOVR views the entire disk of Earth, illuminated near local noon. This provides an ideal geometry for studying the Earth as a proxy exoplanet seen near opposition relative to its parent star. Since the observations are all near full-phase configuration, similar to those near secondary eclipse (when the planet is blocked by the star), the phase-angle effect (Jiang et al. 2018) is not considered in this work. The EPIC instrument images the Earth every 68–110 minutes, returning a total of 9740 frames over a two-year period (2016–2017), with a 2048×2048 charge-coupled device (CCD) in 10 narrowband channels (317, 325, 340, 388, 443, 552, 680, 688, 764 and 779nm), which are selected primarily for investigations of the Earth's climate. A sample observation of reflectance in the 680nm channel at 9:27 UTC, 2017 February 8th is shown in Figure 4.1a.

We integrate the spatially-resolved images over the Earth's disk to simulate observations of an exoplanet that is detected as a point source. This results in a mean reflectance of 0.22 for the image shown in Figure 4.1a. At each time step, reflectance images obtained sequentially from the 10 channels, with exposure time ranging from 22 ms to 654 ms, are combined to form a 10-point reflection spectrum. Since the surface materials and cloud distributions on an Earth-like exoplanet may be significantly different from those on Earth, we do not assume any known spectral features of the surface; instead, we label each reflection spectrum with disk-averaged fractions of land and clouds on the sunlit face of the Earth to evaluate the cause of the change in Earth's light curves. The two fractions are computed as weighted averages, where the weights are proportional to the cosine of the solar zenith angle. Figure 4.1b shows

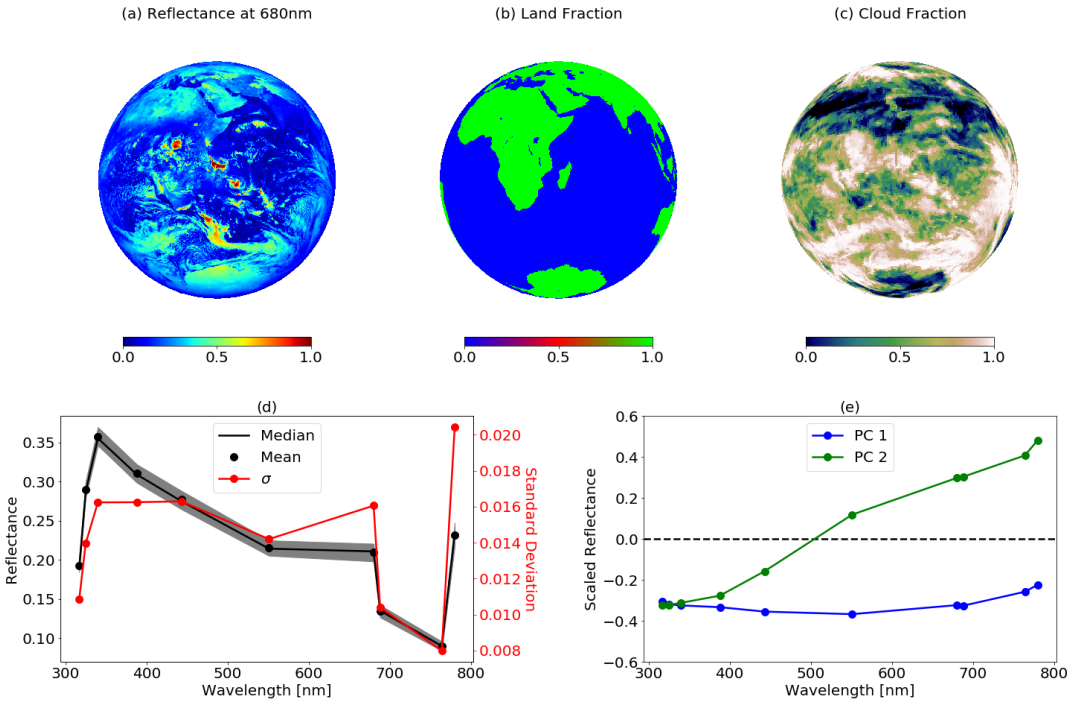

Figure 4.1. (a) Reflectance image in the 680nm channel of DSCOVR/EPIC obtained at 9:27 UTC, 2017 February 8th. The average reflectance is 0.22. (b) Land/ocean map of the Earth for the same scenario as (a), using the GSHHG database (Wessel et al. 1996). The average land fraction is 0.33. (c) Cloud fraction map for the same scenario as (a), obtained from the Level-3 MODIS Atmosphere Daily Global Product (Platnick et al. 2015). The averaged cloud fraction is 0.61. (d) Median (black solid line), mean (black dots) and standard deviation (red line and dots) of reflectance in each channel for ~10,000 DSCOVR/EPIC observations during the years 2016 and 2017. The grey shaded area shows the first and third quartiles of reflection spectra. (e) The first two principal components, PC1 (blue) and PC2 (green), of the scaled reflection spectrum time series.

a land/ocean map as seen from the same viewing angle, with a land fraction of 0.33 as defined by the Global Self-consistent, Hierarchical, High-resolution Geography Database (Wessel et al. 1996). We use the Level-3 MODIS Atmosphere Daily Global Product (Platnick et al. 2015) to compute the cloud fraction label at each time step. Cloud fractions are linearly interpolated between days, since there are multiple EPIC observations each day. This results in a 0.61 cloud fraction for the example observation shown in Figure 4.1c.

In summary, using EPIC observations collected during 2016 and 2017 and the concomitant viewing geometry, we obtain a time series consisting of 10-point reflection spectra with two

coverage fraction labels, which are combined in Section 4.4 to analyze the time series of light curves and to recover the surface map of the proxy exoplanet. The mean and variance of the reflectance in each channel are shown in Figure 4.1d; clearly, there is considerable difference among the channels. In the following analysis, we normalize the time series of reflection spectra to yield zero means and unit standard deviations, since the singular value decomposition (SVD), presented in Section 4.4, is sensitive to the scaling between dimensions of the dataset. This is done in order to give each channel equal importance, since the wavelengths are likely to be different for future exoplanet observations.

4.4 Time series analysis

Time series of disk-averaged light signals carry information about the spectral variability of exoplanets. Analyzing such signals from Earth (in this case, serving as a proxy exoplanet) could provide a baseline for future exoplanet studies. Jiang et al. (2018) used the same dataset and analyzed irradiances from individual EPIC channels to correlate their changes with different types of reflective surfaces, using the original high-resolution images and known spectra of materials on Earth. In their analysis, Jiang et al. (2018) qualitatively explained the variations of single-point light curves for the years 2016 and 2017. However, for future exoplanet studies, Earth-like planets will only be resolved as disk-integrated point sources. Therefore, methods are needed to retrieve information about exoplanet environments from these disk-integrated, point-source observations. For Earth, the light curve is dominated primarily by the cloud cover and land/ocean fraction. Here, we adopt these two parameters, viz., the land and cloud fraction labels, as metrics for evaluating the success of our analysis technique. Therefore, all types of land surfaces are considered the same and an "averaged" land is used in the analysis. Spectra of any reflective surfaces are assumed to be unknown since those on exoplanets could be very different.

We use SVD to decompose the time series into principal components (PCs), and then separate the influences of different reflective surfaces. The first two PCs of the scaled light curves, or "eigencolors" as defined by Cowan et al. (2009), are shown in Figure 4.1e. Given that the singular values are the square roots of the variance along corresponding dimensions,

the first two PCs contain 96.2% of the variation of the scaled light curves (Figure 4.2a). The ordering of the variance accounted for by the PCs depends on the contribution of the land and cloud fractions to the time series variance. Although the PCs are orthogonal, changes in land and cloud fraction are not independent from each other. Land and ocean are fixed to Earth's surface, and appear periodically in the time series of the light curve. Clouds are variable; some parts of Earth are perennially covered by clouds (e.g., the southern oceans), while some others are always clear (e.g., the Sahara Desert). Therefore, some clouds may be synchronized with the rotation of Earth's surface into and out of the instrument field of view, precluding the separation of these clouds from surface features based solely on single-point observations. Nevertheless, this information is encoded in the time series of the disk-integrated image. Information about the perennially cloud or perennially clear scenarios (hereafter referred to as surface-correlated clouds) is expected to be included in the same PC for land/ocean, while that of the rest of clouds would be in another PC.

In order to interpret the physical meaning of the PCs, it is necessary to analyze the relationship between the two labels (land and cloud fractions) and the reflectance time series. We use a machine learning method, called Gradient Boosted Regression Trees (GBRT;

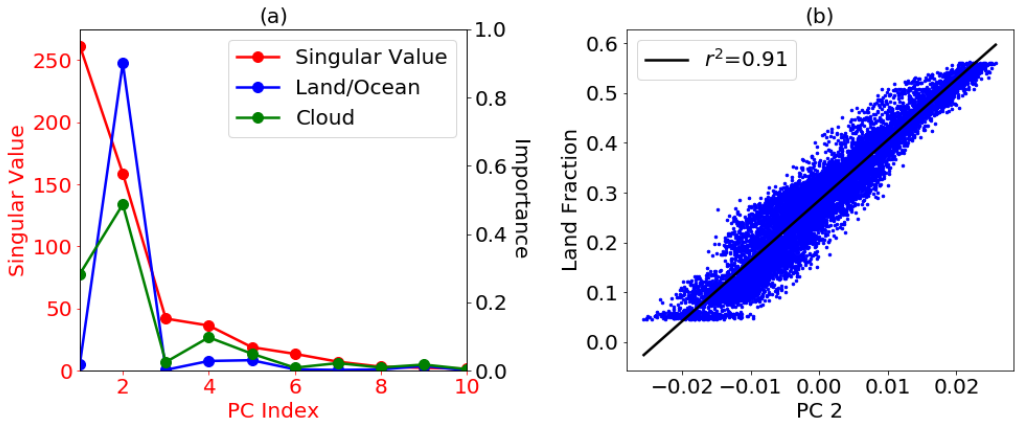

Figure 4.2. (a) Singular values of the principal components (PCs, red) and their importance to land (blue) and cloud (green) fractions. The importance of PCs for each fraction is evaluated using a Gradient Boosted Regression Trees model. (b) Scatter plot of the second principal component, PC2, as a function of land fractions (blue). The best fit line is shown in black, with a correlation coefficient of r^2=0.91.

Friedman 2001), to evaluate the importance of the PCs for each label (Figure 4.2a). This technique computes the relative importance of each PC and denotes them with weights, where the sum of the weights is normalized to unity. Further details on implementing the GBRT technique are presented in Appendix 4.7.1. Clearly, the first two PCs (hereafter, PC1 and PC2) are the most critical, while the contributions from other PCs are negligible. They also have considerably different weights for the land and cloud fractions. For the land fraction, the weight of PC2 is 0.88, while that of PC1 is only 0.03 (Figure 4.2a). Changes in the land/ocean fraction are independent of PC1 and mostly correlated with PC2 (r^2=0.91; Figure 4.2b). For the cloud fraction, the two PCs have comparable weights, 0.28 for PC1 and 0.49 for PC2 (Figure 4.2a). Given the strong correlation between PC2 and surface features, the comparable importance of PC1 and PC2 for clouds suggests that the clouds consist of two types: surface-independent clouds and surface-correlated clouds. This confirms the conclusions of our qualitative analysis above that some changes in clouds can correlate temporally with the surfaces underneath. The importance of clouds in PC2 is likely to be due to these surface-correlated clouds. Conversely, PC1 contributes the largest variation to the

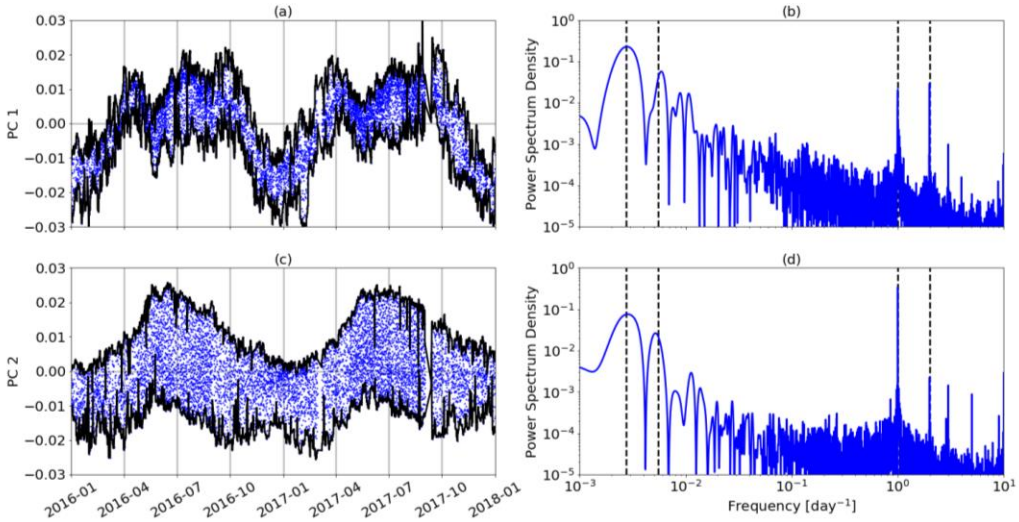

Figure 4.3. (a) Time series of the first principal component, PC1 (blue points). The envelops of daily maxima and minima are denoted by black lines. (b) Power spectrum of the time series of PC1. Cycles of annual, semiannual, diurnal, and half-daily are denoted as black dashed lines. (c) and (d) are identical to (a) and (b), respectively, but correspond to PC2.

light curves via surface-independent clouds that are not correlated with the land/ocean fraction.

The interpretation of the first two PCs is also supported by their time series (Figure 4.3). The time series for PC1 (Figure 4.3a) has more scatter than that for PC2 (Figure 4.3c). The envelope of daily maximum and minimum for PC2 changes gradually between consecutive days, while that for PC1 changes drastically, which indicates that PC2 is more likely to represent features that are surface-associated than detached. Moreover, PC2 peaks in summer, when the northern hemisphere is facing the sun and the spacecraft. The change of PC2 within a day is constant throughout most of the observations, due to the diurnal cycle and the Earth's longitudinal land fraction asymmetry. Fourier analysis shows that both PC1 and PC2 have annual, semi-annual, diurnal, and half-daily cycles (Figures 4.3b and 4.3d). The diurnal cycle of PC2 has the strongest signal in the power spectrum, while that of PC1 is relatively weak. This may be due to the fact that land/ocean reappear with small changes between two consecutive days, while surface-independent clouds can be significantly different in the same time period.

Although convoluted, information on the spatial distribution of different types of surfaces and clouds is fully contained in the time series of an observed planet's light curves. As discussed above, we separate the clouds from surface features using SVD. Surface information about the Earth is mostly contained in PC2 with a strong linear correlation. Here we report the first 2D surface map of Earth (Figure 4.4a) that is reconstructed from single-point light curves using the following assumptions. For the purpose of retrieving the map, the viewing geometry is assumed to be known in this work and obtained from DSCOVR navigation data based on maneuvers that took place during the two-year observation period. It can, in principle, also be derived using light curves and other data (e.g., radial velocity, transit timing) as discussed in more detail in Section 4.5. In the construction of Earth's surface map, spectral features of reflective surfaces are assumed to be unknown in order to facilitate generalization for future Earth-like exoplanet observations. We make the minimal

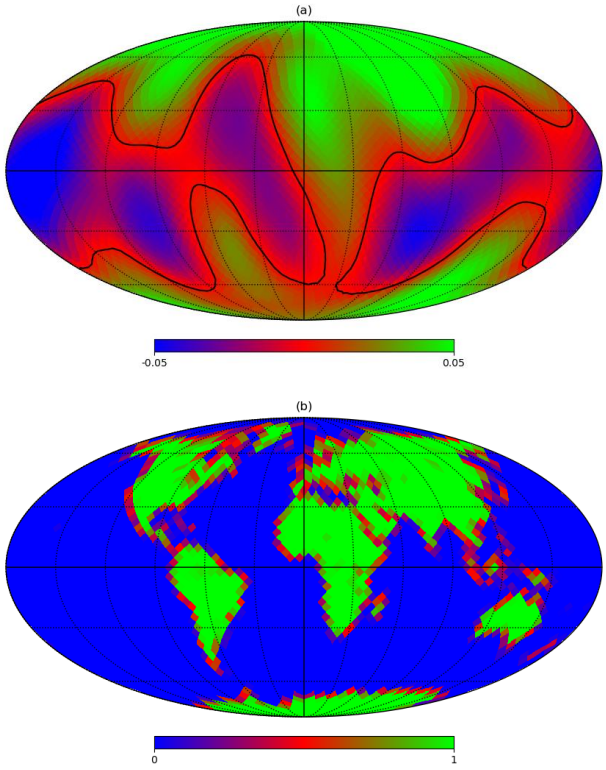

Figure 4.4. (a) 2D surface map of the Earth, treated as a proxy exoplanet, constructed using the PC2 time series. The contour of the median value is given by the black line, which serves as the coastline. The regularization parameter, λ, is 10^{-3} for constructing this map (see Appendix 3.7.2 for further details). (b) Global land/ocean map of the Earth.

assumptions that the incoming solar flux is uniform and known, and that the entire surface of the proxy exoplanet acts as a Lambertian reflector. Although Earth's ocean is strongly non-Lambertian, we still employ the Lambertian assumption because we assume that the surface properties are unknown. This may overestimate the ocean contribution at large distances from the specular point (glint spot) and underestimate it at the specular point. With these assumptions, constructing the Earth's surface map becomes a linear regression problem. Mathematical details are provided in Appendix 4.7.2, and uncertainty estimation is discussed in Appendix 4.7.3. We set the regularization parameter to be 10^{-3} for producing the optimal surface map; results for other values are given in Appendix 4.7.2. The quantity derived in the map is the value of PC2, which has a positive linear correlation with the land fraction as noted above. Coastlines in the reconstructed map are determined by the median value of PC2, which is consistent with the minimal assumption of the overall land fraction

being unknown. Compared with the true land/ocean map (Figure 4.4b), the retrieved map successfully recovers all of the major continents, while there exist some disagreements over oceans. This may be due to the fact that there is often significant cloud coverage over oceans, which reduces sensitivity to surface information in the observations.

4.5 Discussion

A critical requirement for constructing a 2D surface map for an exoplanet is the assumption that the surface information can be extracted from light curves. In the case of Earth, acting as a proxy exoplanet, SVD of light curves can successfully separate Earth's surface from surface-independent clouds. However, the relationships between PCs and features may not be the same for Earth-like exoplanets. This will depend on whether the surface type or clouds introduce measurable variations in the light curves. For Earth, the surface-independent clouds contribute 70.3% of the total variance of scaled light curves; the contribution of the surface is 25.9% (Figure 4.2a). On an exoplanet with less cloud cover, the ratio of these two numbers can be different or even less than unity, which would result in a switch between the first two PCs. In extreme cases where an exoplanet is either cloudless or fully covered by clouds, there will be only one dominant PC instead of the two comparable ones found in this analysis. The third or fourth PC may also be comparable if there exist changes that are significant in spatial and/or spectral scales, e.g. large-scale hydrological processes on continents or another layer of clouds with different composition. Since there will be no ground truth for an exoplanet, spectral analysis may distinguish the PCs between surface and clouds using appropriate assumptions about atmospheric and surface compositions. This can also be addressed by evaluating their time series. The PC associated with surface-independent features tends to have a more chaotic pattern in its time series (Figure 4.3a), while that associated with the surface is more likely to be periodic (Figure 4.3c). It is also worth noting that the surface features of an exoplanet, which are contained in one of the two PCs, will not necessarily be land and ocean. As long as two different surface types have a large albedo contrast and are non-uniformly distributed around the globe, one of the PCs would contain the changes. Materials that are detached from the surface, such as the surface-

independent clouds in the case of the Earth, would also appear in one of the PCs if they have a large influence on the light curve. If these materials can be constrained by the reflection spectra, their fractions can be derived from the magnitude of the corresponding PC.

Once the surface information is extracted from light curves, the surface map of the exoplanet can be recovered from the observational geometry without making any spectral assumptions. Besides orbital elements, which can be determined from light curve observations, the only two geometry assumptions required for constructing the 2D surface map are the summer/winter solstice and the obliquity. The rotation period of a clear or partially-cloudy exoplanet can be inferred from the power spectrum of PCs using a Fourier transform (Figure 4.3d), which requires the observation frequency to be higher than that of the exoplanet's rotation. Studies have been performed to identify a planet's rotational period from light curves at different viewing geometries (e.g. Pallé et al. 2008). The summer/winter solstice would coincide with the maxima and minima of the PC time series as long as the asymmetry between the northern and southern hemispheres is noticeable, and the reflection changes monotonically with latitude when the sub-stellar point is near extremum. For the Earth, the peak of the time series of PC2 takes place on June 15th, 2016 (Figure 4.3c), which is within one week of the true value. The obliquity of an exoplanet could be derived through its influence on the light curves. A number of recent publications (e.g. Schwartz et al. 2016; Kawahara 2016) developed methods for deriving the obliquity using its influence on the amplitude and frequency of light curves. Although these inversion methods are mostly based on a cloudless Earth and known surface spectral features, our SVD analysis of separating clouds from the surface could fill the gap.

Some issues exist in constructing and interpreting the retrieved Earth surface map. Degeneracy resulting from the convolution between pixel geometry and spectrum is the dominant factor affecting map construction quality. As discussed in Cowan & Strait (2013) and Fujii et al. (2017), light curves only cover a small portion of the PC plane, which results in a tradeoff between the spatial and spectral variation. Although only PC2 is used for

constructing the surface map in this work, its time series covers only a small range of all valid values, whose corresponding land fractions are between 0 and 1. Therefore, we introduce a regularization parameter, λ, when constructing the map, to constrain pixel values of PC2 within a reasonable range; the resulting effect is described in Appendix 4.7.2. We select the value of λ based on the ground truth of Earth's surface map. Glints, features that are small in area but contribute significantly to the spectrum, may also influence the quality of the retrieved map. Their contributions to light curves are simulated and estimated in Lustig-Yaeger et al. (2018), and observational evidence in DSCOVR/EPIC images are reported by Li et al. (2018). In this work, the effect of glint is assumed to be on the "average ocean" and removed when scaling the light curves.

4.6 Summary

Spectrally-dependent, single-point light curves of the Earth were analyzed as observations of a proxy exoplanet. SVD analysis suggests that the majority of the information is captured by two principal components. The first captures the non-periodic behavior of surface-independent clouds. The second describes more periodic surface albedo structure. Using the fact that SVD separates the clouds from the surface, we derive the first 2D surface map of the Earth, acting as a proxy exoplanet, from single-point light curves, assuming only that the surface acts as a Lambertian reflector. The geometry is assumed to be known in the analysis, but in principle, it can be derived directly from light curves. This study serves as a baseline for analyzing observations of Earth-like exoplanets with unknown surfaces and possible clouds, enabling future assessments of habitability.

4.7 Appendix

4.7.1 Gradient Boosted Regression Trees

The decision tree model is a decision support tool, which is widely used in the field of machine learning. It uses a tree structure to classify data and make predictions. An example of decision trees used in this work is shown in Figure 4.S1. At each parent node, data points

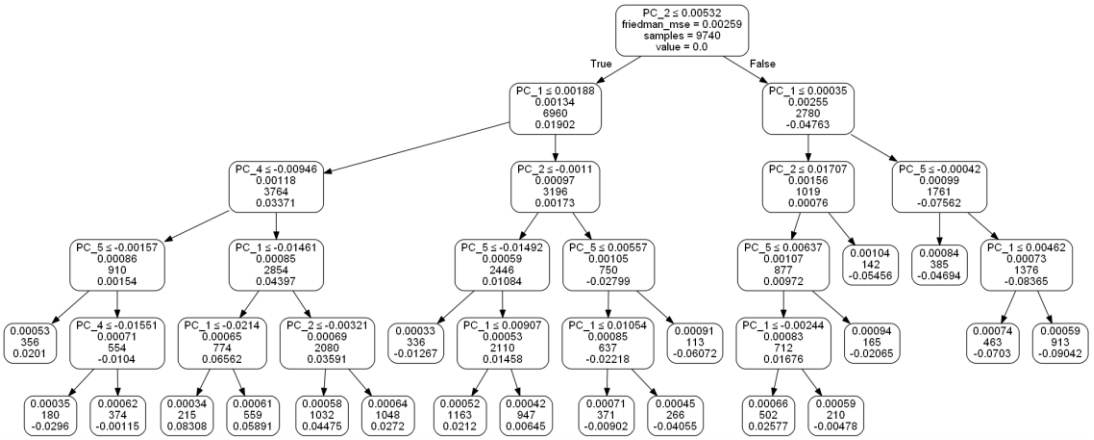

Figure 4.S1. The first decision tree in the Gradient Boosted Regression Trees (GBRT) model. The text in each node shows the criterion, the mean square error (MSE), the number of samples and the averaged label value of the node, respectively. The leaf nodes (nodes that do not have child nodes) only have the latter three.

are divided into two groups, called child nodes, by introducing a threshold for one of the PCs. The label of each child node, cloud fraction in this case, is computed as the mean value of data labels in this group. The mean squared error (MSE) between the prediction and individual labels in each node is evaluated accordingly. Given the finite number of probable ways to divide a parent node, there exists a best PC and a best threshold so that the average MSE of its two child nodes, weighted by their sample numbers, is minimized. Therefore, a decision tree can be constructed using this criterion from a root node, which contains all data points, and the labels are offset in order to have zero mean. For regularization, we introduce a maximum tree depth, a maximum total node number, and a minimum node size to avoid over-fitting.

A machine learning technique, gradient boosting (Friedman 2001), is deployed to improve the model performance and reduce bias. Gradient boosting is an ensemble method that combines weak prediction models, shallow decision trees in this case, to make a final decision. The first decision tree is constructed using the original data points; starting from the second one, the decision tree fits the residual left by the previous tree. Therefore, the final decision is made as the sum of all decision trees. Since weak models tend to have large bias and small variance, while complex models have large variance and small bias, the boosting

technique strikes a balance to achieve the minimum error. We use a Python package developed for machine learning, scikit-learn (Pedregosa et al. 2011), to develop this GBRT model. DSCOVR observations from the years 2016 and 2017 serve as training data, while those from the year 2018 are used as test data. The MSE of the test data is used to select the regularization parameters. The final GBRT model has 250 shallow decision trees, and each decision tree has a maximum depth of 5, a maximum total node number of 20 and a minimum node size of 100. One of the advantages of decision trees is that they can evaluate the feature importance (PCs in this work). We compute the importance of each PC as the number of times they appear as the threshold in the decision tree nodes, weighted by the number of node samples, and normalized to have a unit sum. Results of using land and cloud fractions as the labels are shown in Figure 4.2a. PC2 shows dominant correlation to the land fraction, while PC1 and PC2 show comparable correlation for clouds.

4.7.2 Surface map construction

Given the assumptions discussed in Section 4.5, the averaged reflectance in the *i-th* channel, R_i, can be parameterized as follows:

$$R_i = \sum_p w_p r_{i,p} \qquad (4.S1)$$

where $r_{i,p}$ is the reflectance of the *p-th* pixel of an arbitrary map at the *i-th* wavelength; w_p is the weight of the *p-th* pixel, which is determined by the viewing geometry and has the following form:

$$w_p = \begin{cases} c\cos(\alpha_p)\cos(\beta_p) & \text{when } \alpha_p < 90° \text{ and } \beta_p < 90° \\ 0 & \text{otherwise} \end{cases} \qquad (4.S2)$$

where α_p, β_p are the solar and the spacecraft zenith angles of the *p-th* pixel, respectively; c is a normalization term such that the weights, w_p, sum to unity. The spacecraft is on a halo orbit around L1, which introduces differences between α_p and β_p. The sub-spacecraft point can have a solar zenith angle as large as ~7 °at some time points. Due to the linearity of scaling and SVD, the averaged PC2 at each time point has the same form as the reflectance.

The observed time series of PC2, v_t is given by:

$$v_t = \sum_p w_{t,p} x_p \qquad (4.S3)$$

where x_p is the value of PC2 at the *p-th* pixel in the retrieved Earth map; $w_{t,p}$ has the same form as w_p except that it varies with time. Including the entire time series, this becomes a linear regression problem:

$$\mathbf{W}_{[T\times P]}\mathbf{X}_{[P\times 1]} = \mathbf{V}_{[T\times 1]} \qquad (4.S4)$$

where **W**, **X** and **V** are the matrices with elements $w_{t,p}$, x_p and v_t, respectively. A penalty term, the L^2-norm of **X**, is added to the squared error as a regularization parameter to prevent x_p from deviating too far from zero and therefore avoid over-fitting. The regularized square error, *e*, can be expressed as:

$$e = |\mathbf{WX} - \mathbf{V}|_2^2 + \lambda|\mathbf{X}|_2^2 \qquad (4.S5)$$

where λ is the regularization parameter. Consequently, the solution to minimizing e is as follows:

$$\mathbf{X} = (\mathbf{W}^T\mathbf{W} + \lambda\mathbf{I})^{-1}\mathbf{W}^T\mathbf{V} \qquad (4.S6)$$

where **I** is the identity matrix.

We use the Hierarchical Equal Area isoLatitude Pixelization method (HEALPix; Górski et al. 2005) to pixelate the retrieved map. This technique divides the Earth's surface into pixels with the same area and distributed uniformly on the sphere, appropriate for the DSCOVR observing geometry. The parameter N_{side} in HEALPix is set to 16, which results in a 3072-pixel map with a spatial resolution of ~4°. After solving the regularized linear regression problem, we construct the first 2D surface map of Earth (Figure 4.4a) from single-point light curves.

The parameter λ is selected using synthetic data, where elements in **V** are replaced by time

series of the land fraction label. Three recovered maps using synthetic data are shown in Figure 4.S2, where λ has values of 10^{-4}, 10^{-3} and 10^{-2}, respectively. Comparing them with the ground truth, a value of 10^{-3} is seen to be optimal for λ. Due to the degeneracy in recovering maps from single-point observations and imperfect geometry assumptions, which includes unequal pixel weights and pixelization approximations, the map cannot be further improved even with perfect spectral observation. Fortunately, the map is not sensitive to λ near its optimal value; changing it by one order of magnitude only results in small changes in the coastlines (Figures 4.S2a and 4.S2c). Therefore, we propose that 10^{-3} is a good choice for the regularization parameter for Earth-like exoplanets if observations have comparable numbers of pixels and time steps. The value should be adjusted according to the ratio of the two terms on the right hand side of Equation (4.S5) when the number of pixels and/or time steps are different.

When the land fraction is not known, the selection of λ becomes arbitrary, with the only constraint being that the resulting range of land fraction should be physically valid under the given assumptions. Two more possible maps reconstructed using time series of PC2 with different values of λ (10^{-4} and 10^{-2}) are shown in Figure 4.S3. Comparing the maps recovered using the observations and known ground truth (Figures 4.4a, 4.S3, and 4.S2), it is evident that clouds over oceans contribute considerably to the differences, since the land fraction label is not affected by clouds.

4.7.3 Surface map uncertainty

We estimate the uncertainty in the retrieved Earth surface map (Figure 4.4a) in this section. The observation uncertainty is neglected, since at each time step ~10^6 pixels are averaged so that the observational uncertainty is reduced by a factor of ~10^3. Therefore, we mainly focus on the uncertainty in the linear regression presented in Appendix 4.7.2.

We rewrite Equation (4.S4) with a vector $\mathbf{U}$ for the "true values" of PC2 at each pixel as:

$$\mathbf{W}_{[\mathrm{T}\times\mathrm{P}]}\mathbf{U}_{[\mathrm{P}\times 1]} + \boldsymbol{\varepsilon}_{[\mathrm{T}\times 1]} = \mathbf{V}_{[\mathrm{T}\times 1]} \qquad (4.S7)$$

where ε represents the noise at each time step, and is assumed to follow a Gaussian distribution, $\mathcal{N}(0, \sigma^2\mathbf{I}_{[T*T]})$. An unbiased estimate of σ^2 can be obtained as follows:

$$\sigma^2 = \frac{(\mathbf{V}-\mathbf{W}\mathbf{X})^{\mathrm{T}}(\mathbf{V}-\mathbf{W}\mathbf{X})}{\mathrm{T}-\mathrm{P}} \qquad (4.S8)$$

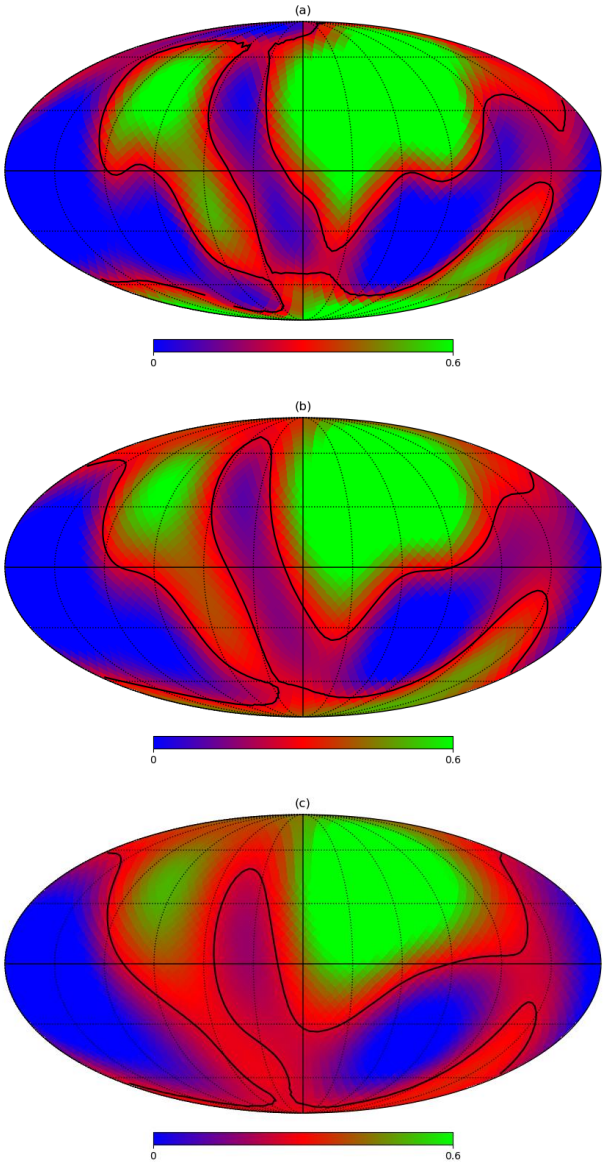

Figure 4.S2. (a) Recovered land fraction map using synthetic observations, produced by averaging the ground truth of the land/ocean map given the viewing geometry. The contour of the median value is given by the black line. The regularization parameter, λ, is 10^{-4} for constructing this map. (b), (c): Same as (a), but for $\lambda=10^{-3}$ and 10^{-2}, respectively.

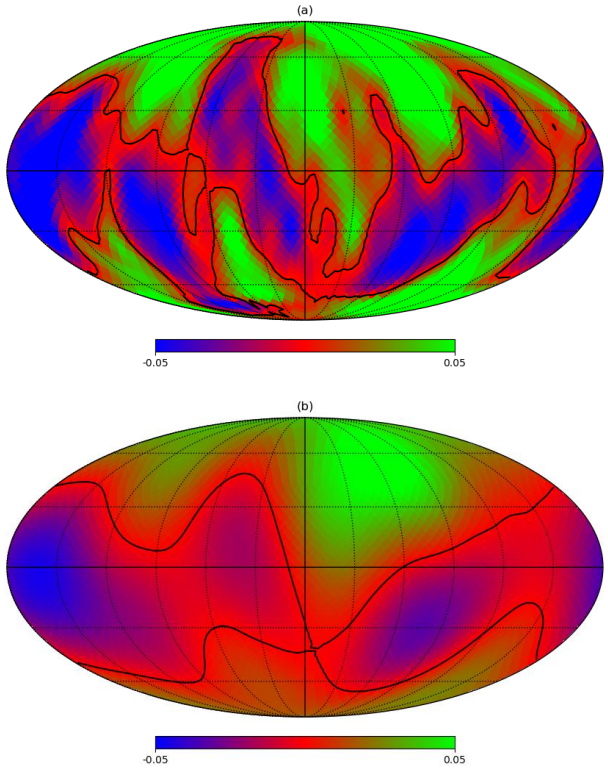

Figure 4.S3. (a), (b): Same as Figure 4.4a, but for $\lambda=10^{-4}$ and 10^{-2}, respectively.

where T and P are the total numbers of time steps and pixels. The difference between these two quantities is the degree of freedom. Combining Equations (4.S6) and (4.S7), $\mathbf{X}$ becomes a Gaussian vector:

$$\mathbf{X} = (\mathbf{W}^{\mathrm{T}}\mathbf{W} + \lambda\mathbf{I})^{-1}(\mathbf{W}^{\mathrm{T}}\mathbf{W})\mathbf{U} + (\mathbf{W}^{\mathrm{T}}\mathbf{W} + \lambda\mathbf{I})^{-1}\mathbf{W}^{\mathrm{T}}\boldsymbol{\varepsilon} \qquad (4.S9)$$

Then, the expectation and covariance matrices of $\mathbf{X}$ can be derived as follows:

$$\mathbf{E}[\mathbf{X}] = (\mathbf{W}^{\mathrm{T}}\mathbf{W} + \lambda\mathbf{I})^{-1}(\mathbf{W}^{\mathrm{T}}\mathbf{W})\mathbf{U} \qquad (4.S10)$$

$$\mathbf{Cov}[\mathbf{X}] = (\mathbf{X} - \mathbf{E}[\mathbf{X}])(\mathbf{X} - \mathbf{E}[\mathbf{X}])^{\mathrm{T}} = \sigma^2(\mathbf{W}^{\mathrm{T}}\mathbf{W} + \lambda\mathbf{I})^{-1}\mathbf{W}^{\mathrm{T}}\mathbf{W}(\mathbf{W}^{\mathrm{T}}\mathbf{W} + \lambda\mathbf{I})^{-1} \qquad (4.S11)$$

The square root of the diagonal elements in the covariance matrix are the 1-sigma uncertainty values for the retrieved map (Figure 4.S4). The uncertainty map is consistent with the viewing geometry; since the sub-solar and the sub-spacecraft points are always near the

equator, pixels at lower latitudes have higher weights, and the uncertainties increase toward the poles. The uncertainty values are on the order of ~10% of the pixel values in the retrieved map (Figure 4.4a), which suggests a good quality of Earth surface map reconstruction.

Chapter 5

Conclusions and Future Directions

The thesis has addressed a range of topics on characterizing astrobiology-relevant environments from prebiotic atmospheres to a currently inhabited planet. Understanding these environments are important and increasingly urgent given the timelines of near future planetary science missions. Chapters presented in this work have addressed several questions on chemical and physical processes in the prebiotic-like atmospheres, and provide benchmarks for observing potentially habitable exoplanets.

Chapter 2 proposed a new methodology to retrieve vertical profiles of chemical species in Titan's upper atmosphere, which can therefore be used to constrain photochemistry and atmospheric dynamics. During the entire Cassini mission, tens of stellar occultations are available to provide such information. Their time coverage spans from late summer to winter in the southern hemisphere of Titan. Therefore, immediate extension of applying the methodology to all available flybys could provide insights into seasonal changes in Titan's atmosphere. The change of methane vertical profile is a tracer of atmospheric dynamics, which is a result of the variation of circulations in Titan's atmosphere. Also, changes of larger organic molecules can provide indications of seasonal changes of chemistry, which is influenced by both of the solar radiation and transport in the atmosphere. Moreover, the results are also spatial-dependent. They can be used to constrain two dimensional general circulation models (GCMs), which can be further coupled with chemical models. This inspires the development of fully coupled atmospheric models of prebiotic atmospheres. The atmosphere of Titan is currently the only one that could provide such detailed observation constraints.

Chapter 3 addressed a fundamental and important question of the morphology of Pluto's haze, which is critical for studying the Pluto system. The bimodal distribution of haze particles satisfies all available observations obtained by the instruments onboard the New Horizons spacecraft, and it likely remains as the best solution until the next mission to Pluto.

The result inspires modeling of microphysical processes of haze formation in Pluto's atmosphere. The bimodal distribution indicates possible dimension transition and/or gas condensation in the atmosphere. Moreover, the precise description of the haze morphology can greatly improve the current photochemical models where the haze particles serve as the condensation/sticking cores, as well as constraining Pluto's surface color, which has indication of Pluto's composition. Above all, observations of the atmosphere of Pluto provides constraints of a cold version of prebiotic-like atmospheres. It can complement observations of Titan's atmosphere, and could be used to cross validate models of prebiotic atmospheres, which have significance in assessing habitability of prebiotic environments.

Chapter 4 investigated the fundamental question "are we alone?" from the other way. It studied the only confirmed inhabited planet, the Earth, from an exoplanet perspective. The analysis showed that information of the spatial features of Early-like exoplanets are included in their light curves, which in principle can be derived even when the resolution is not better than a single pixel. Some questions still remain that the principal components with indices larger than two may be related to minor spatial features or the atmosphere, which are worth being investigated in following works. The progress of this work serves as a benchmark of characterizing Earth-like exoplanets using multi-wavelength single-point light curve observations, and inspires the real exoplanet observations by near future missions. It also provides a new approach for assessing the similarity between potentially habitable exoplanets and the Earth, and therefore their habitability.

It is the author's intention that the thesis will serve to motive continued investigation to the environments related to the emergence and existence of life. Notably, the characterization of Titan's and Pluto's using new observations calls for modeling work to understand the compelling and urgent scientific question about the chemistry and formation of haze in prebiotic atmosphere and their influence on the potentially detectable biosignatures. Moreover, combining such Early-Earth-like observations with those of the current Earth, a complete picture of the distant appearance of the Earth through geological time can be

obtained. It is a guidance for looking for habitable planets beyond solar system and illuminates the way to find the ultimate answer to the fundamental question “are we alone?”.

BIBLIOGRAPHY

Atreya, S. K., et al. (2006), Titan's methane cycle, *Planetary and Space Science*, 54, 1177–1187, doi:10.1016/j.pss.2006.05.028

Au, J. W., Cooper, G., et al. (1993) The valence shell photoabsorption of the linear alkanes, C_nH_{2n+2} (n=1–8): absolute oscillator strengths (7–220 eV), *Chemical Physics*, 173, 209-239. doi: 10.1016/0301-0104(93)80142-V

Buratti, B.J., et al. (2017). Global albedos of Pluto and Charon from LORRI New Horizons observations, *Icarus*, 287, 207–217. doi: 10.1016/j.icarus.2016.11.012

Campbell, B., et al. (1988). A Search for Substellar Companions to Solar-type Stars, *The Astrophysical Journal*, 331, 902-921. doi: 10.1086/166608

Capalbo, F. J., et al. (2016). New benzene absorption cross sections in the VUV, relevance for Titan's upper atmosphere. *Icarus*, 265, 95-109. doi: 10.1016/j.icarus.2015.10.006

Catling, D. C. & Zahnle, K. J. (2020) The Archean atmosphere, *Science Advances*, 6, eaax1420. doi: 10.1126/sciadv.aax1420

Chen, F. Z. & Wu, C. Y. R. (2004). Temperature-dependent photoabsorption cross sections in the VUV-UV region. I. Methane and ethane. *Journal of Quantitative Spectroscopy and Radiative Transfer*, 85, 195-209. doi: 10.1016/S0022-4073(03)00225-5

Cheng, A. F., et al. (2008). Long-Range Reconnaissance Imager on New Horizons, *Space Science Reviews*, 140, 189–215. doi: 10.1007/s11214-007-9271-6

Cheng, A. F., et al. (2017). Haze in Pluto's atmosphere, *Icarus*, 290, 112–133. doi: 10.1016/j.icarus.2017.02.024

Chiang, R., et al. (1993). *Robust attitude control for Cassini spacecraft flying by Titan*. Paper presented at Guidance, Navigation and Control Conference, Monterey, CA. doi: 10.2514/6.1993-3754

Connors, R. E., et al. (1974) Vacuum ultraviolet spectroscopy of cyanogen and cyanoacetylenes, *The Journal of Chemical Physics*, 60, 5011-5024.

Cowan, N. B. & Fujii, Y. (2018). Mapping Exoplanets, in *Handbook of Exoplanets*. doi: 10.1007/978-3-319-55333-7_147

Cowan, N. B. & Strait, T. E. (2013). Determining Reflectance Spectra of Surfaces and Clouds on Exoplanets, *The Astrophysical Journal*, 765, L17. doi: 10.1088/2041-8205/765/1/L17

Cowan, N. B., et al. (2009). Alien Maps of an Ocean-bearing World, *The Astrophysical Journal*, 700, 915–923. doi: 10.1088/0004-637X/700/2/915

Cowan, N. B., et al. (2011) Rotational Variability of Earth's Polar Regions: Implications for Detecting Snowball Planets, *The Astrophysical Journal*, 731, 76-90. doi: 10.1088/0004-637X/731/1/76

Elliot, J. L., et al. (1989). Pluto's atmosphere. *Icarus*, 77, 148–170. doi: 10.1016/0019-1035(89)90014-6

Esposito, L. W., et al. (2004). The Cassini Ultraviolet Imaging Spectrograph investigation. *Space Science Reviews*, 115, 299-361. doi: 33 10.1007/s11214-004-1455-8

Farr, B., et al. (2018). exocartographer: A Bayesian Framework for Mapping Exoplanets in Reflected Light, *The Astronomical Journal*, 156, 146-153. doi: 10.3847/1538-3881/aad775

Ferradaz, T., et al. (2009). Temperature-dependent photoabsorption cross-sections of cyanoacetylene and diacetylene in the mid- and vacuum-UV: Application to Titan's atmosphere. *Planetary and Space Science*, 57, 10-22. doi: 0.1016/j.pss.2008.10.005

Foreman-Mackey, D., et al. (2013). emcee: The MCMC Hammer, *Publications of the Astronomical Society of the Pacific*, 125, 306-312. doi: 10.1086/670067

Friedman, J. H. (2001). Greedy Function approximation: a gradient boosting machine, *The annals of Statistics*, 29, 1189-1232.

Fujii, Y. & Kawahara, H. (2012). Mapping Earth Analogs from Photometric Variability: Spin-Orbit Tomography for Planets in Inclined Orbits, *The Astrophysical Journal*, 755, 101-114. doi: 10.1088/0004-637X/755/2/101

Gao, P., et al. (2017). Constraints on the microphysics of Pluto's photochemical haze from New Horizons observations, *Icarus*, 287, 116–123. doi: 10.1016/j.icarus.2016.09.030

Ge, H., et al. (2019). Rotational Light Curves of Jupiter from Ultraviolet to Mid-infrared and Implications for Brown Dwarfs and Exoplanets, *The Astronomical Journal*, 157, 89-103. doi: 10.3847/1538-3881/aafba7

Geissler, P., et al. (1995). Galileo multispectral imaging of Earth, *Journal of Geophysical Research*, 100, 16895–16906. doi: 10.1029/95JE01407

Gillon, M., et al. (2017). Seven temperate terrestrial planets around the nearby ultracool dwarf star TRAPPIST-1, *Nature*, 542, 456–460. doi: 10.1038/nature21360

Gladstone, G. R., et al. (2016). The atmosphere of Pluto as observed by New Horizons, *Science*, 351, 1280-1286, 2016. doi: 10.1126/science.aad8866.

Górski, K. M., et al. (2005). HEALPix: A Framework for High-Resolution Discretization and Fast Analysis of Data Distributed on the Sphere, *The Astrophysical Journal*, 622, 759–771. doi: 10.1086/427976

Grundy, W. M., et al. (2018). Pluto's haze as a surface material, *Icarus*, 314, 232–245. doi: 10.1016/j.icarus.2018.05.019

Hanel, R., et al. (1981), Infrared observations of the Saturnian system from Voyager 1, *Science*, 212, 192–200. doi: 10.1126/science.212.4491.192

Hapke, B., Bidirectional reflectance spectroscopy. 1. Theory, *Journal of Geophysical Research*, 86, 4571–4586, doi: 10.1029/JB086iB04p03039.

Hastings, W. K. (1970). Monte Carlo sampling methods using Markov chains and their applications. *Biometrika*, 57, 97-109. doi:10.2307/2334940

Hillier, J., et al. (1990). Voyager Disk-Integrated Photometry of Triton, *Science*, 250, 419–421. doi: 10.1126/science.250.4979.419

Hillier, J., et al. (1991). Voyager photometry of Triton: Haze and surface photometric properties, *Journal of Geophysical Research*, 96, 19203–19209. doi: 10.1029/91JA01736

Hörst, S. M., et al. (2018). Haze production rates in super-Earth and mini-Neptune atmosphere experiments. *Nature Astronomy*, 2, 303-306. doi: 53 10.1038/s41550-018-0397-0

Jiang, J. H., et al. (2018). Using Deep Space Climate Observatory Measurements to Study the Earth as an Exoplanet, *The Astronomical Journal*, 156, 26-42. doi: 10.3847/1538-3881/aac6e2

Kameta, K., et al. (2002). Photoabsorption, photoionization, and neutral-dissociation cross sections of simple hydrocarbons in the vacuum ultraviolet range. *Journal of*

Electron Spectroscopy and Related Phenomena, 123, 225-238. doi: 10.1016/S0368-2048(02)00022-1

Kammer, J. A., et al. (2013). Composition of Titan's upper atmosphere from Cassini UVIS EUV stellar occultations. *Planetary and Space Science*, 88, 86-92. doi: 61 10.1016/j.pss.2013.08.003

Kawahara, H., et al. (2016). Frequency Modulation of Directly Imaged Exoplanets: Geometric Effect as a Probe of Planetary Obliquity, *The Astrophysical Journal*, 822, 112-122. doi: 10.3847/0004-637X/822/2/112

Khare, B. N., et al. (1984). Optical constants of organic tholins produced in a simulated Titanian atmosphere: from soft X-ray to microwave frequencies. *Icarus*, 60, 127-137. doi: 10.1016/0019-1035(84)90142-8

Knutson, H. A., et al. (2007). A map of the day-night contrast of the extrasolar planet HD 189733b, *Nature*, 447, 183–186. doi: 10.1038/nature05782

Koskinen, T. T., et al. (2011). The mesosphere and lower thermosphere of Titan revealed by Cassini/UVIS stellar occultations. *Icarus*, 216, 507-534. doi: 10.1016/j.icarus.2011.09.022

Kunde, V. G., et al. (1981), C_4H_2, HC_3N and C_2N_2 in Titan’s atmosphere, *Nature*, 292, 686–688. doi:10.1038/292686a0

Kutsop, et al. (2020). Pluto’s haze abundance and size distribution from limb scatter observations by MVIC. In review.

Lavvas, P., et al. (2010). Titan’s vertical aerosol structure at the Huygens landing site: Constraints on particle size, density, charge, and refractive index, *Icarus*, 210, 832–842. doi: 10.1016/j.icarus.2010.07.025

Lavvas, P., et al. (2013). Aerosol growth in Titan's ionosphere, *Proceedings of the National Academy of Science*, 110, 2729–2734. doi: 10.1073/pnas.1217059110

Lee, L. C. (2008) CN($A^2\Pi_i \rightarrow X^2\Sigma^+$) and CN($B^2\Sigma^+ \rightarrow X^2\Sigma^+$) yields from HCN photodissociation, *The Journal of Chemical Physics*, 72, 6414-6421. doi: 10.1063/1.439140

Levenberg, K. (1944). A method for the solution of certain non-linear problems in least squares, *Quarterly of Applied Mathematics*, 2, 164-168. doi: 10.1090/qam/10666

Li, C., et al. (2014). A non-monotonic eddy diffusivity profile of Titan's atmosphere revealed by Cassini observations. *Planetary and Space Science*, 104, 48-58. doi: 10.1016/j.pss.2013.10.009

Li, C., et al. (2015). Vertical distribution of C_3-hydrocarbons in the stratosphere of Titan. *The Astrophysical Journal Letters*, 803, L19. doi: 71 10.1088/2041-8205/803/2/L19

Li, J.-Z., et al. (2019). Study of Terrestrial Glints Based on DSCOVR Observations, *Earth and Space Science*, 6, 166–173. doi: 10.1029/2018EA000509

Liang, M. -C., et al. (2007). Photolytically generated aerosols in the mesosphere and thermosphere of Titan. *The Astrophysical Journal Letters*, 661, L199. doi: 10.1086/518785

Lindal, G. F., et al. (1983), The atmosphere of Titan—An analysis of the Voyager 1 radio occultation measurements, *Icarus*, 53, 348–363, doi:10.1016/0019-1035(83)90155-0

Louden, T. & Wheatley, P. J. (2015). Spatially Resolved Eastward Winds and Rotation of HD 189733b, *The Astrophysical Journal Letters*, 814, L24. doi: 10.1088/2041-8205/814/2/L24

Luspay-Kuti, A., et al. (2017). Photochemistry on Pluto - I. Hydrocarbons and aerosols, *Monthly Notices of the Royal Astronomical Society*, 472, 104–117. doi: 10.1093/mnras/stx1362

Lustig-Yaeger, J., et al. (2018). Detecting Ocean Glint on Exoplanets Using Multiphase Mapping, *The Astronomical Journal*, 156, 301-316. doi: 10.3847/1538-3881/aaed3a

Maltagliati, L., et al. (2015). Titan's atmosphere as observed by Cassini/VIMS solar occultations: CH4, CO and evidence for C2H6 absorption. *Icarus*, 248, 1-24. doi: 10.1016/j.icarus.2014.10.004

NASA Exoplanet Archive, Confirmed Planets Table, doi: 10.26133/NEA1

NASA-PDS (2017) *Cassini UVIS User's Guide*, Planetary Data System (http://pds-rings.seti.org/cassini/uvis/index.html)

Nuth, J. A. & Glicker, S. (1982) The vacuum ultraviolet spectra of HCN, C_2N_2, and CH_3CN, *Journal of Quantitative Spectroscopy and Radiative Transfer*, 28, 223-231. doi: 10.1016/0022-4073(82)90025-5

Pallé, E., et al. (2008). Identifying the Rotation Rate and the Presence of Dynamic Weather on Extrasolar Earth-like Planets from Photometric Observations, *The Astrophysical Journal*, 676, 1319–1329. doi: 10.1086/528677

Pantos, E., et al. (1978) The extinction coefficient of benzene vapor in the region 4.6 to 36 eV, *Journal of Molecular Spectroscopy*, 72, 36-43. doi: 10.1016/0022-2852(78)90041-3

Pedregosa F., et al. (2011). Scikit-learn: Machine Learning in Python, *Journal of Machine Learning Research*, 12, 2825-2830.

Pilinski, E. B., & Lee, A. Y. (2009). Pointing-stability performance of the Cassini spacecraft. *Journal of Spacecraft and Rockets*, 46, 1007-1015. doi: 95 10.2514/1.41675

Platnick, S., King, M., & Hubanks, P. 2015. MODIS Atmosphere L3 Daily Product. doi:10.5067/MODIS/MOD08_D3.006; doi:10.5067/MODIS/MYD08_D3.006

Ramirez, S. I., et al. (2002). Complex Refractive Index of Titan's Aerosol Analogues in the 200-900 nm Domain, *Icarus*, 156, 515–529. doi: 10.1006/icar.2001.6783

Rannou, P., et al. (1997). A new interpretation of scattered light measurements at Titan's limb, *Journal of Geophysical Research*, 102, 10997–11014. doi: 10.1029/97JE00719

Reuter, D. C., et al. (2008). Ralph: A Visible/Infrared Imager for the New Horizons Pluto/Kuiper Belt Mission, *Space Science Reviews*, 140, 129–154. doi: 10.1007/s11214-008-9375-7

Sagan, C., et al. (1993). A search for life on Earth from the Galileo spacecraft, *Nature*, 365, 715–721. doi: 10.1038/365715a0

Schwartz, J. C., et al. (2016). Inferring planetary obliquity using rotational and orbital photometry, *Monthly Notices of the Royal Astronomical Society*, 457, 926–938. doi: 10.1093/mnras/stw068

Schwieterman, E. W., et al. (2018). Exoplanet Biosignatures: A Review of Remotely Detectable Signs of Life, *Astrobiology*, 18, 663–708. doi: 10.1089/ast.2017.1729

Seager, S. & Bains, W. (2015) The search for signs of life on exoplanets at the interface of chemistry and planetary science, *Science Advances*, 1, e1500047. doi: 10.1126/sciadv.1500047

Shemansky, D. E., & Liu, X. (2012). Saturn upper atmospheric structure from Cassini EUV and FUV occultations. *Canadian Journal of Physics*, 90, 817-831. doi: 10.1139/p2012-036

Shemansky, D. E., et al. (2005). The Cassini UVIS stellar probe of the Titan atmosphere. *Science*, 308, 978-982. doi: 10.1126/science.1111790

Spurr, R. & Natraj, V. (2011). A linearized two-stream radiative transfer code for fast approximation of multiple-scatter fields, *Journal of Quantitative Spectroscopy and Radiative Transfer*, 112, 2630–2637. doi: 10.1016/j.jqsrt.2011.06.014

Stern, S. A., et al. (2008). ALICE: The Ultraviolet Imaging Spectrograph Aboard the New Horizons Pluto-Kuiper Belt Mission, *Space Science Reviews*, 140, 155–187. doi: 10.1007/s11214-008-9407-3

Stern, S. A., et al. (2015). The Pluto system: Initial results from its exploration by New Horizons, Science, 350. doi: 10.1126/science.aad1815

Tomasko, M. G., et al. (2008). A model of Titan's aerosols based on measurements made inside the atmosphere, *Planetary and Space Science*, 56, 669–707. doi: 10.1016/j.pss.2007.11.019

Vervack, R. J., et al. (2004). New perspectives on Titan's upper atmosphere from a reanalysis of the Voyager 1 UVS solar occultations. *Icarus*, 170, 91-112. doi: 10.1016/j.icarus.2004.03.005

Wessel, P. & Smith, W. H. F. (1996). A global, self-consistent, hierarchical, high-resolution shoreline database, *Journal of Geophysical Research*, 101, 8741–8743. doi: 10.1029/96JB00104

Willacy, K., et al. (2016). A new astrobiological model of the atmosphere of Titan. *The Astrophysical Journal*, 829, 79-89. doi: 10.3847/0004-637X/829/2/79

Wiscombe, W. J., (1979). MIE scattering calculations, advances in technique and fast, vector-shaped computer codes, *NCAR Technical Note*. doi: 10.5065/D6ZP4414

Wong, M. L., et al. (2017). The photochemistry of Pluto's atmosphere as illuminated by New Horizons, *Icarus*, 287, 110–115. doi: 10.1016/j.icarus.2016.09.028

Wu, C. Y. R., et al.(2001). Measurements of temperature-dependent absorption cross sections of C_2H_2 in the VUV-UV region. *Journal of Geophysical Research*, 106, 7629-7638. doi:128 10.1029/2000JE001292

Wu, C. Y. R., et al.(2004). Temperature-dependent photoabsorption cross sections in the VUV-UV region: Ethylene. *Journal of Geophysical Research Planets*, 109, E07S15. doi: 10.1029/2003JE002180

Young, L. A., et al. (2018). Structure and composition of Pluto's atmosphere from the New Horizons solar ultraviolet occultation, *Icarus*, 300, 174–199. doi: 10.1016/j.icarus.2017.09.006

Zhang, X., et al. (2017), Haze heats Pluto’s atmosphere yet explains its cold temperature, *Nature*, 551, 352–355. doi: 10.1038/nature24465